Measures of University Research Output

DISSERTATION

zur Erlangung des akademischen Grades

doctor rerum politicarum
(Doktor der Wirtschaftswissenschaft)

eingereicht an der

Wirtschaftswissenschaftlichen Fakultät
der Humboldt-Universität zu Berlin

von

Alona Zharova

Präsident der Humboldt-Universität zu Berlin:
Prof. Dr.-Ing. Dr. Sabine Kunst

Dekan der Wirtschaftswissenschaftlichen Fakultät:
Prof. Dr. Christian D. Schade

Gutachter:
1. Prof. Dr. Wolfgang Karl Härdle, Humboldt-Universität zu Berlin
2. Prof. Dr. Stefan Lessmann, Humboldt-Universität zu Berlin

Tag des Kolloquiums: 08.02.2018

Acknowledgement

First and foremost, I would like to express my sincere gratitude to my first advisor Wolfgang Karl Härdle for continuous support of my doctoral studies. I am thankful for his inspiring and wise guidance and the creation of a highly supportive environment. His friendly advice during our work together and guidance throughout the academic journey have helped me reach this stage in my life. My sincere thanks also go to my second advisor Stefan Lessmann for his insightful and instructive comments. His supportive feedback motivated me greatly and gave me the incentive to widen my research and to look beyond a limited scope.

This research was financially supported by the German Research Foundation (DFG) via Collaborative Research Center 649 "Economic Risk" (CRC 649) and International Research Training Group 1792 "High Dimensional Nonstationary Time Series" (IRTG 1792), Humboldt-Universität zu Berlin. I am very thankful for the support of the Vice-President for Research, Peter A. Frensch, who granted access to the data and for the critical input of Ingmar Schmidt and Carsten Gerrits from the Research Service Centre of the Humboldt-Universität zu Berlin. The support from Inga Link and her colleagues from Elsevier is gratefully acknowledged. I highly appreciate the valuable insights and suggestions of Bernd Fitzenberger. I would like to express special thanks to Jeffrey Wooldridge for our fruitful discussions, funny stories and good coffee.

I would like to thank the co-authors of my research papers Andrija Mihoci and Janine Tellinger-Rice. Andrija's guidance and encouragement helped me through the difficult initial stages of research. I am deeply thankful to Marius Sterling for his excellent research assistance and enthusiastic support. I am also grateful for the assistance provided by Nicole Hermann und Dominik Prugger.

This thesis was written parallel to my position as a Managing Director of the CRC 649 and later the IRTG 1792. It is a pleasure to express appreciation to my team, who always showed understanding and accomplished their tasks 120%. My thanks go to Raphael Reule, Alisa Kolesnikova, Mona Schirmer, Hilko Paschke, Marcel Mucha, Alexander Meyer and many others. I wish much success you all!

A very special word of gratitude goes to all my friends and colleagues at the Ladislaus von Bortkiewicz Chair of Statistics for their support and input; Cathy Chen, Alla Petukhina, Awdesch Melzer, Hien Pham Thu, Lenka Zbonakova, Ya Qian and many others. In particular, I would like to thank to Leslie Udvarhelyi. His careful English proof reading and fine eye for detail have improved this thesis. Special thanks go to Petra Burdejova for her helpful comments on this thesis, but also for her optimistic attitude and a lot of happy moments. I also thank Rainer Voß for the great technical support.

I give my sincerest thanks to Jan, the dearest person in my life, for his love, patience and continuous support throughout the doctoral phase. His wise counsel, regular inspiring and encouraging words influenced the prompt conclusion of the thesis.

This thesis is dedicated to my parents Olga and Victor for their unconditional support throughout my life and over lasting encouragement through the process of researching and writing this thesis.

Berlin, January 2018 *Alona Zharova*

Abstract

New Public Management helps universities and research institutions to perform in a highly competitive research environment. Decision making in the face of uncertainty, for example distribution of funds for research needs and purposes, urges research policy makers and university managers to understand the relationships between the dimensions of research performance and the resulting or incoming grants. Thus, it is important to accurately reflect the variables of scientific knowledge production on the level of individuals, research groups and universities.

Chapter 2 of this thesis introduces an analysis on the level of individuals. The data are taken from the three widely-used ranking systems in the economic and business sciences among German-speaking countries: Handelsblatt (HB), Research Papers in Economics (RePEc, here RP) and Google Scholar (GS). It addresses the problem that often hampers decision making in academic institutions – incomplete research profiles. It proposes a framework for collating ranking data for comparison purposes.

Chapter 3 provides empirical evidence on the level of research groups using data from a Collaborative Research Center (CRC) on financial inputs and research output from 2005 to 2016. First, suitable performance indicators are discussed. Second, main properties of the data are described using visualization techniques. Finally, the time fixed effects panel data model and the fixed effects Poisson model are used to analyze an interdependency between financial inputs and research outputs.

Chapter 4 examines the interdependence structure between third-party expenses (TPE), publications, citations and academic age using university data on individual performance in different scientific areas. A panel vector autoregressive model with exogenous variables (PVARX), impulse response functions and a forecast error variance decomposition help to capture the relationships in the system. In particular, the analysis quantifies the influence of TPE, publications and citations on each other; the reaction of the system to exogenous impulses; and the proportion of variance explained by considered variables. Besides analyzing on the university level, the data is also reviewed for various faculties, revealing differences between scientific areas. The chapter also proposes a visualization of the cooperation between faculties, and research interdisciplinarity via the co-authorship structure among publications. To summarize, the chapter addresses the possible implications for policy and decision making and proposes recommendations for university research management.

Keywords: Research Performance, Scientometrics, Bibliometrics, Collaborative Research Center, Decision Making, Third-Party Funds, PVARX Model, Time Fixed Effects Panel Data Model, Fixed Effects Poisson Model.

Zusammenfassung

New Public Management unterstützt Universitäten und Forschungseinrichtungen dabei, in einem stark wettbewerbsorientierten Forschungsumfeld zu bestehen. Entscheidungen unter Unsicherheit, z.B. die Verteilung von Mitteln für den Forschungsbedarf und Forschungszwecke, erfordert von Politik und Hochschulmanagement, die Beziehungen zwischen den Dimensionen der Forschungsleistung und den resultierenden oder eingehenden Zuschüssen zu verstehen. Hierfür ist es wichtig, die Variablen der wissenschaftlichen Wissensproduktion auf der Ebene von Individuen, Forschungsgruppen und Universitäten zu untersuchen.

Das Kapitel 2 dieser Arbeit analysiert die Ebene der Individuen. Es verwendet die Beobachtungen der Forscherprofile von Handelsblatt (HB), Research Papers in Economics (RePEc, hier RP) und Google Scholar (GS) als meist verbreitete Ranking-Systeme in BWL und VWL im deutschsprachigen Raum. Dieses Kapitel schlägt einen Rahmen vor, in dem durch Daten aus unterschiedlichen Rankings unvollständige Forschungsprofile ergänzt und vergleichbar gemacht werden können.

Das Kapitel 3 liefert eine empirische Evidenz für die Ebene von Forschungsgruppen und verwendet die Daten eines Sonderforschungsbereichs (SFB) zu Finanzinputs und Forschungsoutput von 2005 bis 2016. Das Kapitel beginnt mit der Beschreibung passender Performanzindikatoren, gefolgt von einer innovativen visuellen Datenanalyse. Im Hauptteil des Kapitels untersucht die Arbeit mit Hilfe eines Zeit-Fixed-Effects-Panel-Modells und eines Fixed-Effects-Poisson-Modells den Zusammenhang zwischen finanziellen Inputs und Forschungsoutputs.

Das Kapitel 4 beschäftigt sich mit dem Niveau der Universitäten und untersucht die Interdependenzstruktur zwischen Drittmittelausgaben, Publikationen, Zitationen und akademischem Alter mit Hilfe eines PVARX-Modells, einer Impulsantwort und einer Zerlegung der Prognosefehlervarianz. Die Ergebnisse erlauben Aussagen über den Zusammenhang zwischen den Forschungsleistungsindikatoren für einzelne Fakultäten und bieten eine Reihe von möglichen Erklärungen für Unterschiede verschiedener Wissenschaftsbereiche. Diese Forschung quantifiziert den Einfluss von Drittmittelausgaben, Publikationen und Zitationen auf einander, die Reaktion des Systems auf exogene Impulse und die Menge an Varianz, die durch berücksichtigte Variablen erklärt wird. Die Ergebnisse sind für jeweils Sozial- und Geisteswissenschaften, Lebenswissenschaften sowie Mathematik und Naturwissenschaften zusammengefasst. In diesem Kapitel wird auch eine Visualisierung der Kooperation zwischen Fakultäten und Forschungsinterdisziplinarität über die Koautorenschaft zwischen den Publikationen vorgeschlagen. Abschließend befasst sich das Kapitel mit den möglichen Implikationen für Politik und Entscheidungsfindung und schlägt Empfehlungen für das universitäre Forschungsmanagement vor.

Schlagwörter: Forschungsleistung, Scientometrie, Bibliometrie, Sonderforschungsbereich, Entscheidungsfindung, Drittmittel, PVARX-Modell, Zeit-Fixed-Effects-Modell, Fixed-Effects-Poisson-Modell.

Contents

List of Figures

List of Tables

List of abbreviations

BWL	Business sciences (germ. Betriebwirtschaftslehre)
CRC	Collaborative Research Center
DFG	German Research Foundation (germ. Deutsche Forschungsgemeinschaft)
DP	Discussion paper
FE	Fixed Effects Panel Data Model
FEP	Fixed Effects Poisson Model
FEVD	Forecast error variance decomposition
HB	Handelsblatt
IRF	Impulse response function
JEL	Journal of Economic Literature classification in the economic sciences
NPM	New Public Management
TPF	Third-party funds
TPE	Third-party expenses
PVARX	Panel vector autoregressive model with exogenous variables
RP	Research Papers in Economics (RePEc)
SP	Sub-projects
VWL	Economic sciences (germ. Volkswirtschaftslehre)
WR	German Council of Science and Humanities (germ. Wissenschaftsrat)

1 Introduction

New Public Management (NPM) helps universities and research institutions to perform in a highly competitive research environment. It emerged in the 1980s (Hood 1991) with the goal of improving efficiency and overall performance of public sector institutions by using business management approaches and models. NPM places a strong focus on permanent monitoring and evaluation of performance. Measuring research performance allows an analysis of the structural issues in science. It can thus facilitate the development of a scientific system and strengthen excellence in research.

Decision making in the face of uncertainty, such as the distribution of funds for research needs and purposes, urges research policy makers and university managers to understand the relationships between the dimensions of research performance and the resulting or incoming grants. Support of the effective decision making process requires both qualitative and quantitative information. It is important to accurately reflect the interdependency between input and output variables of scientific knowledge production on the level of individuals, research groups and universities; and also to account for time-delayed effects with the appropriate methodology.

Chapter 2 introduces an analysis on the level of individuals. The data is taken from the three widely-used ranking systems in economic and business sciences among German-speaking countries: Handelsblatt (HB), Research Papers in Economics (RePEc, here RP) and Google Scholar (GS). For the economic discipline, for which the Handelsblatt (HB) ranking system has become the most recognized platform in Germany, a framework for collating ranking data for comparison purposes is suggested. A single HB common score for scholars within the HB community is proposed, as the result of an analysis of the interconnectedness between HB sub-rankings through quantile regression. The cross-ranking dependence analysis of Handelsblatt, Research Papers in Economics and Google Scholar ranking schemes shows that researcher age and field of specialization – mapped onto the JEL classification codes – have a substantial impact on the resulting scores.

Based on the conducted analyses, chapter 2 shows that quantile regression successfully interpolates and estimates the proposed HB common score. Academic rankings data exhibit different correlation structures over the underlying scores of HB, RP and GS, whereas the academic ranking variation has been documented to be quite sensitive to age differences. For example, the rank of both younger and older scientists is changing marginally (increasing) and is becoming more significant than the rank of middle-aged researchers. The scientists specializing in microeconomics (HB), international economics (RP) and general economics (GS) are associated with the respective leading positions. However, researchers from mathematical and quantitative fields occupy high positions across all three ranking systems.

Chapter 3 provides empirical evidence on the level of research groups using time fixed effects panel data model and fixed effects Poisson model. To study the relationship between research outcomes and funding of a Collaborative Research Center (CRC), the number of discussion papers (DPs) are regressed on staff and travel costs using sub-projects' (SP) level data. With the help of year dummy variables, the chapter shows how the pattern of SP productivity changed from 2006 to 2016 after controlling for staff and travel costs. Since the level of spending from the previous year and the preceding number of DPs may influence the current number of DPs, a control for the lagged variables is added. The productivity of each SP may differ due to some heterogeneity or individual effects, such as skills of a principal investigator (PI), average abilities or skills of researchers employed at the SP or a specific behavior of a research field. For instance, working on a publication with one vs. more co-authors, writing in English vs. other languages, or publishing in books vs. articles may affect the research outcomes. Therefore, the possibility of individual SP effects is allowed.

Chapter 4 contributes to a deeper understanding of the interplay between third-party funds (TPF), publications and citations using university data on individual performance in different scientific areas. A distinctive feature of this study is the analysis of individual-level data from a German university, which belongs to the top 10 universities in Germany in terms of external funds acquisition (DFG 2015). A sample of professorships, the complete set of their third-party expenses (TPE), publications, and citations from Scopus, is observed on a yearly basis for the period 2001 to 2015. Additionally, a variable measuring academic age (number of years after Ph.D. degree) is included. This information enables the analysis on a fine level of granularity and provides the possibility to account for time-delayed effects.

Decision and policy making in research management must take into account research field heterogeneity. Given concerns about the feedback and interdependency, a panel vector autoregressive model with exogenous variable (PVARX) is employed (Canova and Ciccarelli 2013, Cavallari and D'Addona 2014). The PVARX model is estimated for each faculty aiming to underline the existing inter-faculty heterogeneity. The resulting impulse response functions (IRF) help to understand the relationship between variables in a VAR context and clarify how a change in one variable affects another variable. For example, one may be wondering to what extent the number of publications will change, if TPE increase by 1%. Since the analysis of such original innovations is rarely the case in work with real data (Tsay 2014), orthogonalized innovations received, using Cholesky decomposition of the white noise covariance matrix, are used. Finally, a forecast error variance decomposition (FEVD) indicates a percentage of the change in the prediction error that is explained by a shock at a four-year time horizon. The last chapter addresses the possible implications for policy and decision making and proposes recommendations for the university research management.

The statistical analysis is performed using R, MATLAB and Stata. The codes (Quantlets) are available on a web-based repository hosting service and collaboration platform GitHub (2018). The technology of Quantlets is provided in QuantNet (2018), Borke and Härdle (2017) and Borke and Härdle (2018).

2 Individuals: Academic Ranking Scales in Economics

Publication in academic and professional journals is a vital aspect of any scientist's career. The number of media outlets and the quality of published research influences decisions on jobs, salary, tenure and so on. Academic ranking scales, particularly in economics, are commonly used for the classification, judgment and evaluation of the scientific depth of individual research. These ranking systems all compete against each other and allow for different disciplinary gravity to be applied. They try to provide a fair platform for the evaluation of research results at universities, research centers and institutes, interdisciplinary groups, etc.

Ranking systems also play a key role in performance comparison and the clarification of individual contribution to the overall ranking of an institution. For instance, decisions made during recruitment processes at German universities (in economic fields) are typically supported by HB rankings, see Schläpfer and Schneider (2010). Furthermore, the distribution of financial resources at universities is often based on performance-related schemes that include achieved research results being taken into consideration, see Oberschelp and Jaeger (2015).

This chapter deals with the performance analysis of researcher' profiles utilizing ranking observations from the most popular ranking systems in the economic and business sciences among German-speaking countries (Germany, Austria and Switzerland): Handelsblatt (HB), Research Papers in Economics (RePEc, here RP) and Google Scholar (GS) databases. The underlying ideas of these rankings and their comparison is discussed in Butz and Wohlrabe (2016), Wohlrabe (2011), Dilger and Müller (2011).

The research questions include: (i) How HB profiles of researchers can be completed based on the available data of the given HB sub-rankings? (ii) How to impute scores and how to predict an academic rank for researchers, who are not already included in a particular HB sub-ranking system? (iii) How strong is the cross-ranking dependence between the score outputs of HB, RP and GS? (iv) Which variables contribute significantly to ranking's dependence and score results?

Quantile regression offers a more detailed modeling framework than ordinary least-squares or least-absolute deviation fitting. The latter methods model the average response; a comprehensive ranking analysis of researchers should instead focus on other data characteristics, such as quantiles in our case. Quantile regression presently receives relatively close attention from the research community, along with the often used, average-response methods in ranking (citation) analysis employed by e.g. Hamermesh (2015). A comprehensive introduction to the quantile regression method is given in Koenker (2005). The rapidly growing literature shows a variety of approaches and applications in statistics and bibliometrics. Birks et al. (2014) use quantile regression with bootstrapped standard errors to predict the median, the 90^{th} and 95^{th} quantiles of the h-index for researchers in the health care field. For example, quantile regression allows: Rauber and Ursprung (2008) to investigate the research productivity of German academic economists over their life cycles; Kelchtermans and Veugelers (2011) to explore the research performance in relation to different sets of productivity drivers; whereas Stegehuis et al. (2015) predict the number of citations in publications. Here, in this study, we employ quantile regression to complete and define the research profiles of scholars.

The proposed approach and the findings of this research can be successfully used in practice (a) by selection committees in recruitment processes at universities (economic fields), (b) as a unique tool in decision making related to the allocation of research funds, (c) for collaborative purposes and grant proposal applications, etc. Our estimated HB common score can finally and confidently be used for a simultaneous comparison of candidates profiles from business (BWL) and economic (VWL) sciences.

This chapter is structured as follows. The description of the analyzed ranking systems and our data sources is presented in Section 2.1. Section 2.2 describes the statistical modelling steps related to data selection and the implementation of the predicting techniques. Section 2.3 discusses the HB, RP and GS comparison results and provides evidence on the impact of age and the research fields on ranking performance. Finally, Section 2.4 concludes.

2.1 Academic Ranking Systems

In this analysis, the terms *ranking*, *rank* and *score* are repeatedly used. Ranking represents the academic system or scale; rank denotes the position of each individual within the ranking; and score denotes the number of points assigned.

2.1.1 Handelsblatt (HB)

The HB ranking provides a list of the most active researchers publishing in business and economics in Germany, Austria and Switzerland and also German-speaking researchers outside of these countries. The rankings were developed by the Konjunkturforschungsstelle (KOF) of the ETH Zürich on behalf of HB and German Association for Social Policy (Verein für Sozialpolitik). For this purpose the publication data from several external databases and the data from the Forschungsmonitoring (2018) are used. The HB ranking system has an established reputation among German-speaking economists since it influences decision making regarding the distribution of funds, recruitment process and performance evaluations at universities, Schläpfer (2011).

Moreover, HB produces and publishes a journal ranking list compiled from selected journals indexed in The American Economic Association's electronic bibliography (EconLit), see Combes and Linnemer (2010). Every journal from the HB list receives a weight of between 0.05 and 1, where a higher weight indicates a higher rank. An individual researcher's rank is generated from the number of weighted publications in relevant journals divided by the number of co-authors.

HB considers two fields: business sciences (BWL) and economics sciences (VWL). Within each field the following sub-rankings can be found: the Researcher Life's Work (LW); Current Researchers (CR); and Researchers Under 40 (U40). This gives a total of six BWL and VWL sub-rankings that are usually published every 24 months. The CR ranking is based on researchers' publications in predetermined journals over the last five years, whereas the U40 ranking considers all scientists younger than 40. The LW ranking, finally, takes all rated publications from the HB journals' list into account. It is worth noting that each researcher is present in either the VWL or in the BWL ranking, although inside each category, the individual can belong to any of the sub-ranking categories, LW, CR or U40 (the last only if he/she is younger than 40).

Here we utilize the sub-rankings of 250 individuals from VWL LW in 2015 and 250 individuals from BWL LW in 2014. For the sake of brevity, we provide a detailed descriptive analysis with programming codes in GitHub (2018); the results are available from the author upon request. In order to implement the analyses of the research fields and the age of the researchers based on the score, we have had to eliminate the individuals with missing observations, i.e. with no information on age or research fields.

2.1.2 Research Papers in Economics (RP)

The RP ranking system collects the bibliographic data of journal articles, books, working papers and other scientific media outlets. It contains around 2.3 million research items from more than 2,800 journals and 4,500 working paper series, see RePEc (2018). Although the RP project offers a broad spectrum of services, in this paper we focus solely on author ranking. The main idea of the RP author's ranking system is to publish a list of the top 5% researchers on a monthly basis, from a pool of 50,000 registered individuals, based on an average rank score. This score is calculated based on a two-step procedure for each author. First, the authors are individually ranked within each of the 36 separate sub-rankings, excluding the w-index, a special case of the h-index. Second, a harmonic mean of the individual ranks represents this average rank score. In contrast to HB and GS, one should note that within the RP system the top-ranked scientists receive the lowest score and vice versa. For more details, we refer to Zimmermann (2013) and the corresponding RP webpage.

Contrary to HB, all RP sub-rankings receive the same weight while providing the average rank score, although they may impose a weighting scheme. To boost an HB score, for instance, an author must consider the journal ranking list, whereas to improve their RP score, researchers must consider other publication aspects, such as number of citations, abstract views, etc. Since the HB ranks were collected up to 2015 inclusive, the RP data for 2304 individuals were collected for December 2015 (see Table A.1 in Appendix).

2.1.3 Google Scholar (GS)

Contrary to HB and RP, GS concentrates on citation data (Hamermesh 2015). For every researcher, GS provides information about the number of citations per paper, the total number of citations, and the values of the h-index and the $i10$-index. The latest three indicators are here analyzed for 1,438 researchers. While calculating its metrics, GS takes into account all types of research publications. GS has good coverage in social sciences, economics, finance and business administration, see Harzing and Wal (2008), which makes it a desirable choice for our research purposes.

2.1.4 Data

Our work considers HB (2014, 2015), RP (December 2015) and GS data (December 2015). In order to take into account both economic and business sciences, we select two main HB rankings with data available for 500 scientists: (i) the VWL LW in 2015 for 250 individuals and (ii) BWL LW in 2014 for 250 individuals. In December 2015, 2,304 researchers were listed in RP top 5% author ranking. Of those, 1,027 had a GS profile with corresponding GS scores.

A more detailed view of the data merging results is depicted in the mosaic plot, Figure 2.1. Consider the 500 scientists in HB. There are 122 individuals that also have an RP score, but not a GS profile. Similarly, 260 individuals have HB and GS scores, but no RP ranking data. Finally, there are 84 researchers (76 VWL, 8 BWL) for which the HB, RP and GS data are all available.

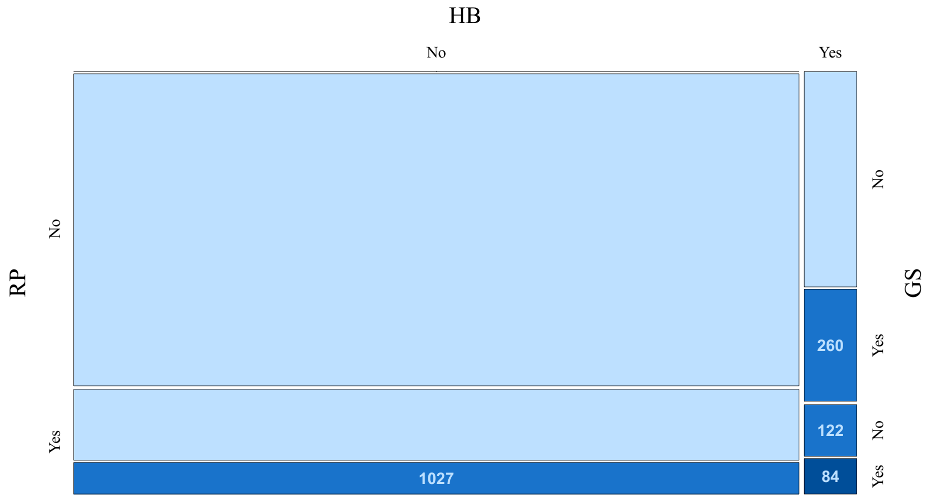

Figure 2.1: Mosaic plot for the number of researchers, whether merging of HB, RP and GS rankings takes place or not (Yes/No). The number of GS profiles is quite large and here they are only shown as an approximation.

2.2 Methodology

Quantile regression offers a more comprehensive description of the relationship between two variables than a linear regression model. A linear regression model considers the relation between the dependent variable and one or more regressors as an average through the conditional mean function. On the contrary, quantile regression offers a broader perspective, since it models various conditional quantile functions, providing the possibility to depict the interconnections at various points, see Koenker (2015) and Baum (2013). For instance, for $\tau = 0.5$ the conditional median function results in a functional that is of limited influence, i.e. robust with respect to outliers. The analysis of data with thick tails and/or non-normal errors may not only turn out to be challenging but may also be biased for the linear model.

2.2.1 Quantile Regression

A linear regression (LR) model

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i,\ i = 1, \ldots, n, \tag{2.1}$$

where β_0 denotes the intercept and β_1 depicts the regression line slope with ε_i denoting the error term models the mean response of variable Y in relation to the regressor X. Here n stands for the sample size, i.e. in our case the number of data (ranking score) pairs $\{y_i, x_i\}_{i=1}^n$. As proposed by Koenker and Bassett (1978) and Koenker and Hallock (2001), we use the quantile regression (QR) model related to the linear regression (2.1) as

$$y_i = \beta_{0,\tau} + \beta_{1,\tau} x_i + \varepsilon_i,\ i = 1, \ldots, n, \tag{2.2}$$

where $\tau \in (0, 1)$ denotes the quantile level and the error ε_i has τ-quantile zero. For instance, setting $\tau = 0.5$ results in median quantile regression.

In the estimation of the linear regression model, the estimates of the unknown intercept and the slope parameter are found by least square minimization

$$\left(\widehat{\beta}_0, \widehat{\beta}_1\right) = \arg\min_{\beta_0,\beta_1} \sum_{i=1}^{n} (y_i - \beta_0 - \beta_1 x_i)^2 \tag{2.3}$$

whereas in quantile regression by the minimization of the asymmetrically weighted residuals

$$\left(\widehat{\beta}_{0,\tau}, \widehat{\beta}_{1,\tau}\right) = \arg\min_{\beta_0,\beta_1} \sum_{i=1}^{n} \rho_\tau\left(y_i - \beta_0 - \beta_1 x_i\right), \tag{2.4}$$

with check function $\rho_\tau(u) = u\{\tau - \mathbf{1}(u < 0)\}$, where $\mathbf{1}(\cdot)$ denotes the indicator function.

2.2.2 HB Common Score

As a practical application of quantile regression for completing of research profiles, our study considers the prediction of HB sub-ranking scores. As there are more VWL researchers (76 individuals) relative to BWL (8 individuals) within the merged dataset (see Figure 2.11), we found it convenient to consider the score of a VWL researcher as the dependent variable and the score of the BWL researcher as the explanatory variable. The resulting *HB common score*, thus, represents the observed and the predicted VWL scores. Consider the 250 VWL LW (y_i), as well as the 250 BWL LW (x_i) scores and then fit the (median) quantile regression (2.4). Denote the estimated model parameters by $\widehat{\beta}_{0,0.5}$ and $\widehat{\beta}_{1,0.5}$. Then the estimated HB common scores for the BWL researchers, using the analysed $n = 250$ pairs (y_i, x_i), are found by

$$\widehat{y}_i = \widehat{\beta}_{0,0.5} + \widehat{\beta}_{1,0.5} x_i, \quad i = 1, \ldots, 250. \tag{2.5}$$

Empirical results show an excellent explanatory performance, see e.g. the scatterplot with imposed fitted median quantile regression line and the Quantile-Quantile (QQ) plot in Figure 2.2, the estimated parameters in Table 2.1, and the goodness-of-fit measures in Table 2.2. The proposed HB common score is represented either by the existing VWL LW score for the VWL researchers or by the predicted score for the BWL researchers. In total, 500 HB common scores are associated with the 500 researchers.

		Est.	SE	t	p-value
BWL LW	$\widehat{\beta}_{1,0.5}$	-0.28	0.21	-1.37	0.1725
	$\widehat{\beta}_{0,0.5}$	1.07	0.04	27.71	0.0000

Table 2.1: Estimated regression model parameters (Est.) for rankings between VWL LW (dependent variable) and BWL LW (explanatory variable) for HB researchers. We provide the standard error of estimates (SE), the t-statistics to test whether the null hypothesis' the true parameter equals 0', and also the associated p-value.

2.2.3 Statistical Analysis

Outliers and extreme values may affect the regression estimation results. Here we first illustrate the robustness of quantile (median) regression to the presence of extreme values as compared with the ordinary least squares regression.

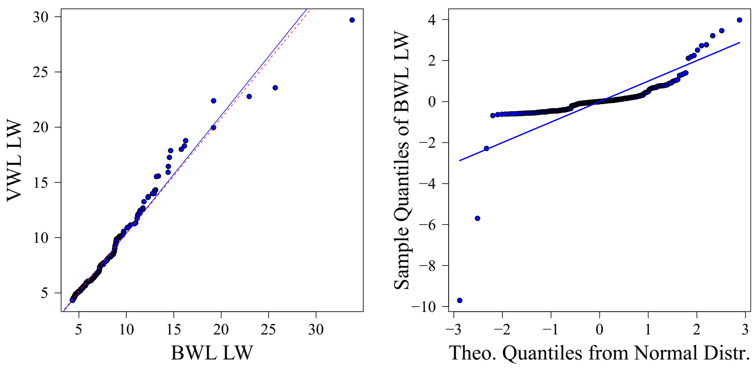

Figure 2.2: Scatterplot and quantile regression fit (left) of the HB on VWL LW vs BWL LW for a sample of 250 researchers within these rankings. Superimposed on the plot is the 0.50 quantile regression line (solid blue) and the least squares estimate of the conditional mean function (dashed red line). The coefficient of determination of the median regression equals 0.93. On the right, a QQ plot of the same sample of data versus a normal distribution.

	MSE	r^2
BWL LW	0.9976	0.9308

Table 2.2: Mean squared error (MSE) and coefficient of determination of the regression model for rankings between VWL LW (dependent variable) and BWL LW (explanatory variable) for HB researchers.

We then study the structural HB score dependence and provide evidence for ranking prediction while changing the underlying quantile level.

In our modelling framework we now consider the data matrix excluding k (largest) observations. For convenience, we select $k \in \{1, 2, 5, 10, 15\}$ and present the resulting parameter estimates for the quantile (median) and linear regression in Table 2.3.

	$k = 0$	$k = 1$	$k = 2$	$k = 5$	$k = 10$	$k = 15$
$\widehat{\beta}_0$	-0.09	-0.50	-0.74	-0.91	-0.72	-0.57
$\widehat{\beta}_1$	1.05	1.10	1.14	1.17	1.14	1.11
$\widehat{\beta}_{0,0.5}$	-0.28	-0.54	-0.59	-0.63	-0.42	-0.21
$\widehat{\beta}_{1,0.5}$	1.07	1.12	1.12	1.13	1.09	1.05

Table 2.3: Estimated parameters using least squares and quantile regression ($\tau = 0.50$) for datasets excluding k largest observations/outliers.

One observes that the estimated quantile regression parameters are more insensitive to the presence of outliers. A relatively lower parameter estimates variability favours the quantile regression as compared to least squares fitting. In practice, our proposed ranking imputation framework is thus a preferable choice.

The presented framework provides an insight into the tail dependence structure of the HB score distribution. In this aspect we consider various quantile levels, namely

$$\tau = \{0.05, 0.25, 0.50, 0.75, 0.95\}.$$

Based on the ranking (BWL) data, one can estimate the corresponding quantiles of the other (VWL) observations, see the results of the employed quantile regression models in Figure 2.3. For example, consider a (top) rated BWL scientist with score 20. The predicted 95^{th} quantile VWL score is near 24, whereas the estimated 5^{th} quantile is close to 18.

Summarising these statistical findings, our ranking imputation approach offers a framework that successfully accounts for the presence of extreme values

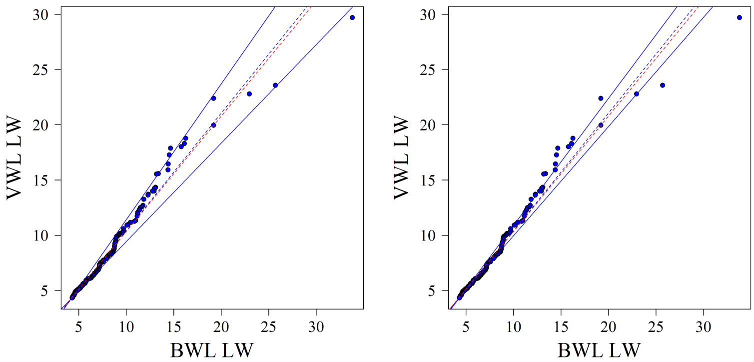

Figure 2.3: Scatterplot and Quantile Regression Fit of the HB on VWL LW vs. BWL LW for a sample of 250 researchers within these rankings. Superimposed on the plots is the 0.05 and 0.95 (left) as well as 0.25 and 0.75 (right) quantile regression line as solid blue, the 0.50 median quantile regression line (dashed blue line) and the least squares estimate of the conditional mean function (dashed red line).

and more importantly, provides valuable results of the score distribution properties. We recommend employing the approach in the recruitment process at universities that consider HB (top) ranking performances.

2.3 Cross-Rankings Dependence

The HB common score is used here in the dependence analysis. First, we show the connection and similarities between the considered rankings; then we investigate the influence of age on the ranking scores. Finally, we provide a detailed analysis of the scores relative to the research fields. Note that here we use HB, RP or GS to denote the HB common score, the RP average rank score and the number of GS citations, respectively.

2.3.1 HB, RP and GS

The distributions of HB and GS scores of researchers are asymmetric, right-skewed and single-peaked, see Figure 2.4. The heavy tails stretching away from the peaks indicate the presence of many outliers that fall outside of the overall pattern, here associated with extreme values. We have a concentration of data in the left part and a long tail to the right. This represents the vast majority of scientists with lower rankings, with only a few individuals possessing very high rankings. In the RP scores distribution, in contrast, one can identify multiple peaks close together. The structure of the RP average rank score can explaine this, as it is calculated from 36 sub-rankings.

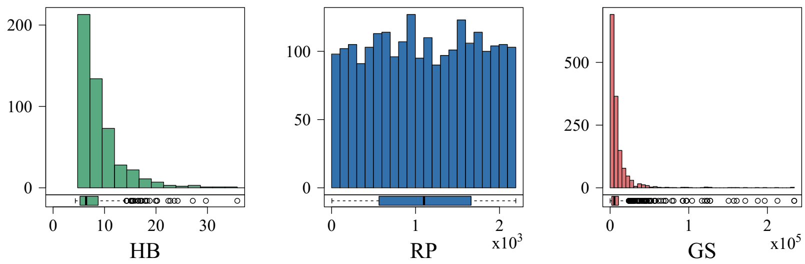

Figure 2.4: Histogram of HB (500 observations, common score), RP (2,304, total score $\times 10^3$) and GS (1,357, citations $\times 10^5$) rankings for December 2015.

One observes a moderate and positive dependence between the HB, RP and GS scores; please see the parallel coordinates plot, Figure 2.5. The three quartiles (25%, 50% and 75%) indicate a considerable number of outliers that influence the results. This can be confirmed by removing the extreme scores from HB and GS. The result is shown in Figure A.1 in the Appendix.

The relationship between HB, RP and GS scores is further analyzed for the full data frame consisting of 42 factors in the correlation matrix in Figure 2.6. Here we use the HB common score and also include the age of researchers as an additional factor. The descriptive statistics is introduced in Table A.1 in the Appendix.

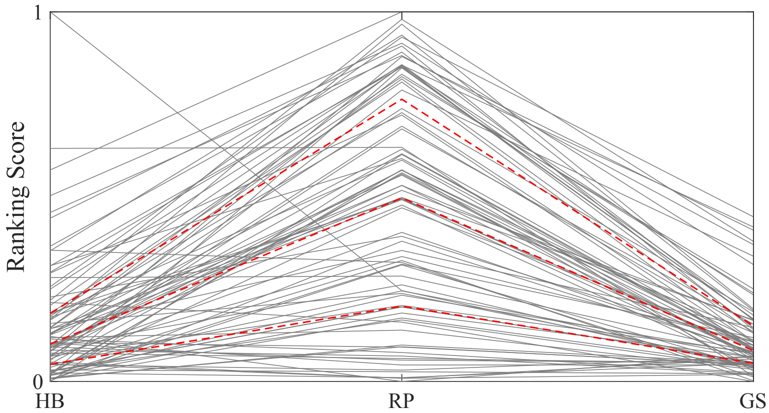

Figure 2.5: Parallel coordinate plot for three variables (HB, RP and GS) on 84 researchers for December 2015. For convenience, the RP values are reversed. Red lines denote the three quartiles (25%, 50% and 75%).

The correlation plot reveals that many variables indicate a strong linear relationship. In particular, the correlation between GS citations and other variables varies, mainly moderate to strong. The HB common score shows, in most cases, a moderate correlation. The visible clusters that characterize RP data correspond to the groups of RP sub-rankings. The negative correlation between RP average rank and other variables is due to the difference in scales, as explained in Section 2.1.2.

One can notice that the RP and GS citations and h-index show a very strong correlation. These pairwise relations are additionally explored through the hexagon plot in Figure 2.7. The Figures indicate a positive linear relationship between the two. However, some outliers that do not follow this trend.

2.3.2 Influence of Age

Our research question is to study whether age influences the rankings of scientists. As the age data is available for only 458 individuals from HB, we have

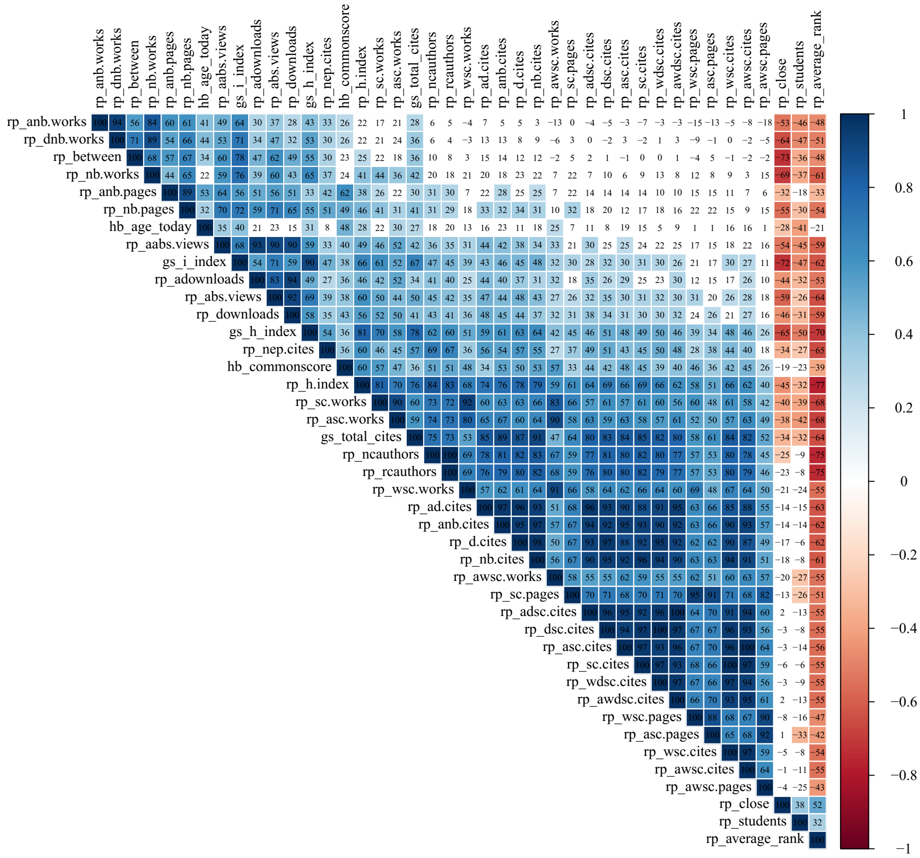

Figure 2.6: Correlation matrix of 42 factors of HB, RP and GS for 84 researchers in December 2015. The color depicts the strength of correlation: from positive (blue) to negative (red).

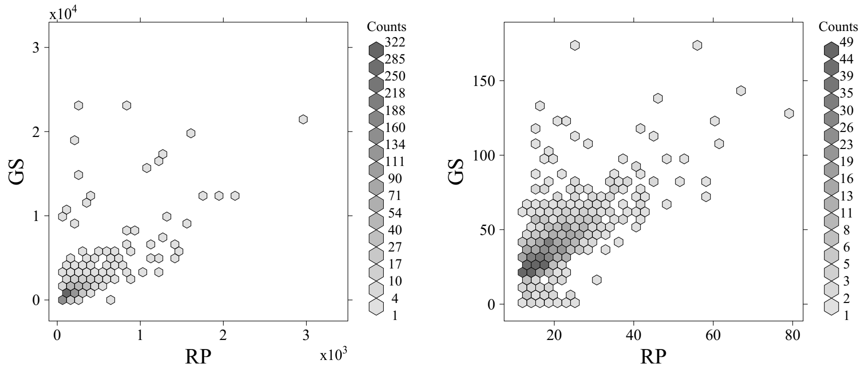

Figure 2.7: Hexagon plot of RP and GS citations for 1024 researchers (left) and hexagon plot of RP and GS *h*-index for 928 researchers (right) in December 2015. Correlation coefficient equals to 0.70 for citations and 0.68 for h-index.

also reduced the full datasets within RP and GS to the top 458 observations. The scatterplots, hexagon plots and boxplots in Figures 2.8 – 2.10 show the relationships between age and ranking scores in a more detailed way.

From Figure 2.8 one can make several observations. Firstly, that a positive relationship between age and HB ranks exists; for RP it is difficult to identify any pattern of data points. Here it is important to note that some RP rankings are standardized with respect to age, while simultaneously there seems to be a very weak association between age and GS.

For the research aggregate, we divide the ranking scores of scientists into nine groups with respect to their age with five-year steps, starting with individuals younger than 36 years and concluding with ones older than 70. The overall patterns of response for the age groups are described on the boxplots in Figures 2.9 – 2.10.

The notable high box length of ranks from the RP age groups indicates of the high sample variability. On the other hand, the comparatively short boxplots from the GS age groups indicate that GS researchers have only slight difference on the introduced scale. In the same way, the boxplots of HB are comparatively tall. This suggests that 458 of the HB scientists have relatively

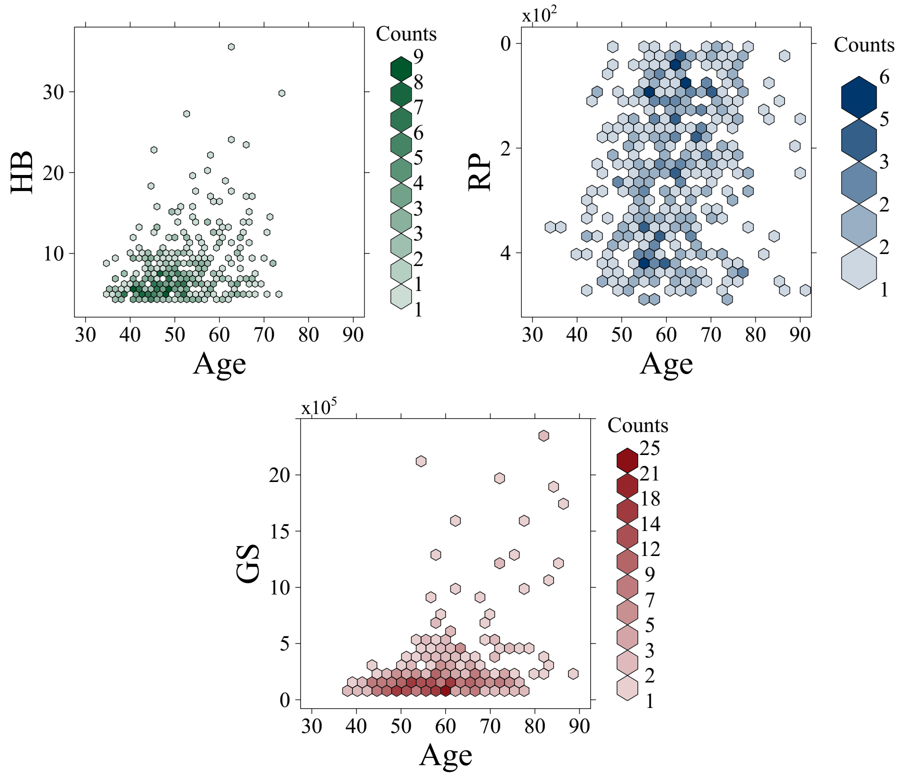

Figure 2.8: Hexagon plots for age and ranking scores of HB, RP and GS for 458 individuals within each ranking system for December 2015.

different ranking scores. Almost all age groups of HB, moreover, indicate the presence of heavy tails in the direction of higher ranks, as, in some cases, the length of the whiskers exceeds the length of the boxes.

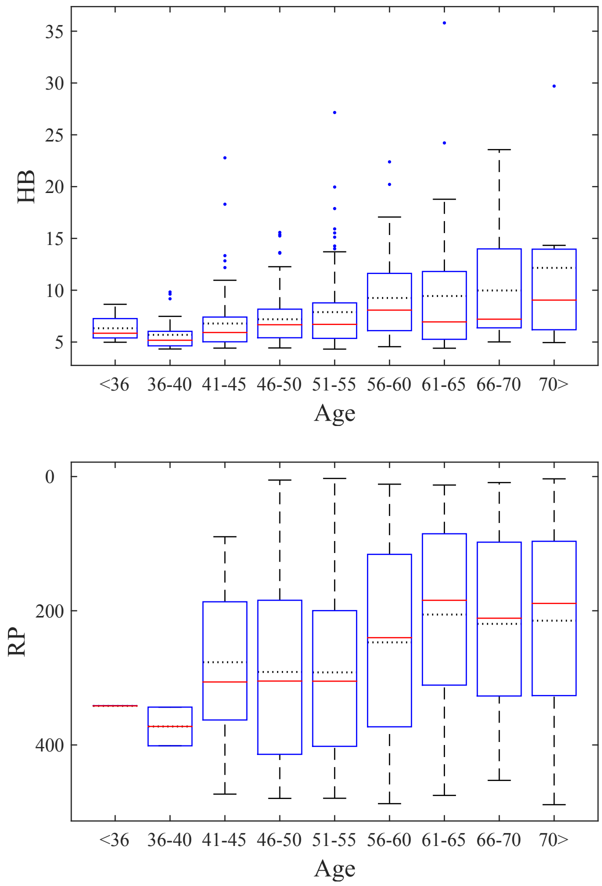

Figure 2.9: Boxplots for age and ranking scores of HB (top) and RP (bottom) for 458 individuals within each ranking system for December 2015. The red lines denote the median, whereas the dotted lines introduce the mean. For comparison purposes the RP scale is inverted.

A further analysis shows that ranks of younger researchers are increasing, whereas the middle-aged group has relative consistency or a slight decline and then the next growth trend, amongst scientists of advanced age, could be observed. One possible explanation for this observation could originate from a

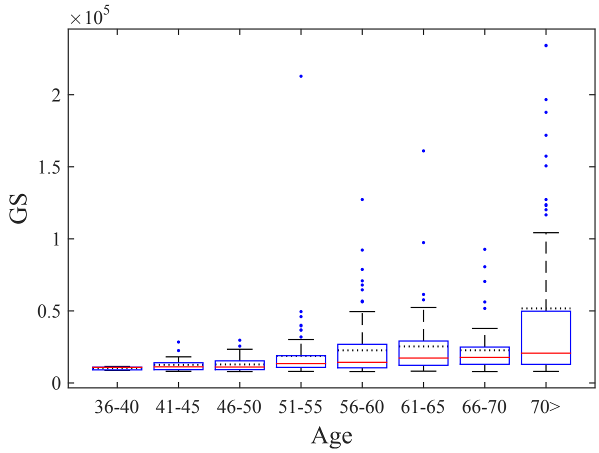

Figure 2.10: Boxplots for age and ranking scores of GS for top 458 individuals within each ranking system for December 2015. The red lines denote the median, whereas the dotted lines introduce the mean.

scientific path in academia. In order to get a position at a university, young researchers are encouraged and motivated to write as many papers as possible and produce other significant outputs, while the middle-aged researchers, who usually have stable positions, concentrate more on teaching, long-term projects and other duties. The slight increase in ranking scores of older individuals could be explained by experience in writing papers, acknowledgment amongst the scientific community, enlarged research networks that they work within and other variables. As a result, this leads to a higher level of work, citations, indexes and number of papers downloaded.

The relative comparison of three academic rankings through the age groups in a four-dimensional plot (HB, RP and GS scores and age) is represented by a mosaic plot in Figure 2.11. We consider three academic rankings with the 458 researchers from each one. Here HB, RP and GS scores are shown by green, blue and red colours respectively. The width of each column represents the number of individuals within each age group, whereas the coloured dot represents zero.

This plot shows that the majority of younger extraordinary researchers belong to the HB group. Amongst the middle-aged ones, the slight domination of

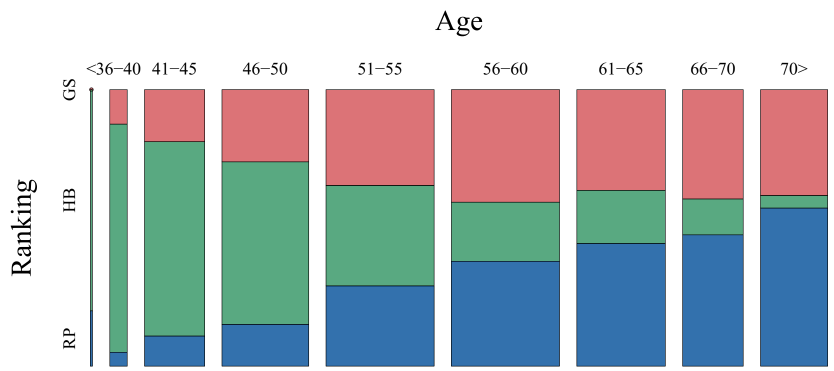

Figure 2.11: Mosaic plot of HB (green), RP (blue) and GS (red) scores for top 458 individuals within each ranking system for December 2015. The width of the columns represents the number of individuals within each age group.

GS over the RP system is visible. At the same time, the scientists of advanced age are mostly located in RP and partly in GS areas.

2.3.3 Research Fields

We were able to enrich our dataset and perform a comparative analysis by adding the research field of scientists provided by HB and GS. From 500 researchers in HB, only 448 individuals have information about subject fields. This constraint forces us to reduce the GS dataset by taking the 448 best ones from Figure 2.1, thus enabling the comparison. From RP we also select the top 448 individuals, although they are from merged GS and HB data; see Figure 2.1. As a result, the RP scientists that originally had no information relating to their areas of research receive these from their GS profiles or their HB ranking systems. Therefore, we end up with a dataset that contains 448 scientists within each of the discussed ranking systems with their main research field.

In order to analyse the influence of research area on ranking scores, all researchers were divided into 19 groups of subject fields according to their recognition classification in economic sciences Journal of Economic Literature (JEL), see JEL (2018). The explanation of the JEL codes is given in Table 2.4.

Code	Research field
A	General Economics and Teaching
B	History of Economic Thought, Methodology, and Heterodox Approaches
C	Mathematical and Quantitative Methods
D	Microeconomics
E	Macroeconomics and Monetary Economics
F	International Economics
G	Financial Economics
H	Public Economics
I	Health, Education, and Welfare
J	Labor and Demographic Economics
K	Law and Economics
L	Industrial Organization
M	Business Administration and Business Economics / Marketing / Accounting / Personnel Economics
N	Economic History
O	Economic Development, Innovation, Technological Change, and Growth
P	Economic Systems
Q	Agricultural and Natural Resource Economics / Environmental and Ecological Economics
R	Urban, Rural, Regional, Real Estate, and Transportation Economics
Y	Miscellaneous Categories
Z	Other Special Topics

Table 2.4: JEL Classification System.

A distribution of scores of researchers within research areas (JEL codes) and the corresponding ranking systems can be seen on the comparative histograms in Figure 2.12. The frequency Table 2.5, generated from our dataset, shows that more than 16% of selected HB researches come from microeconomics (D). They are followed by scientists from the business field (M), financial economics (G), mathematical and quantitative methods (C), and macroeconomics and monetary economics (E), with over 10% within each research area.

	A	B	C	D	E	F	G	H	I	J
GS	86	3	53	32	43	67	46	13	5	22
HB	1	2	49	73	49	39	59	1	6	10
RP	72	2	50	41	68	73	42	14	4	26
	K	L	M	N	O	P	Q	R	Z	
GS	0	13	12	0	29	0	13	9	2	
HB	3	48	67	1	24	0	8	4	4	
RP	1	13	2	1	22	1	7	6	3	

Table 2.5: Frequency Table for JEL codes and the ranking scores of HB, RP and GS for the top 448 scientists within each ranking system for December 2015.

A distinct difference is introduced by RP, where international economics (F), and general economics and teaching (A) hold the leading positions with over 16% for each. Macroeconomics and monetary economics accompany these, along with mathematical and quantitative methods with over 15% and 11% respectively. In the same manner, the dominant research area of GS is presented by general economics and teaching with more than 19% of researchers. Furthermore, international economics produces above 14% of GS, while mathematical and quantitative methods, and financial economics, make up 11% and 10% respectively. However, the mathematical and quantitative methods field is the only research field amongst the ones compared that has over 10% across all three ranking systems.

To sum up, we present a mosaic plot in Figure 2.13 that gives us the advantage of a relative simultaneous comparison of ranking systems through the subject fields in a four-dimensional space. The width of the columns, illustrating the aggregated number of individuals within each research area, brings us to the following important conclusions. Since the F column is a widest one, the largest number of researchers occupying the leading positions among HB, RP and GS carry out their research in international economics. Fields such as macroeconomics and monetary economics, general economics and teaching, mathematical and quantitative methods, microeconomics and financial economics illustrate a slight little difference. On the other hand, the presence of

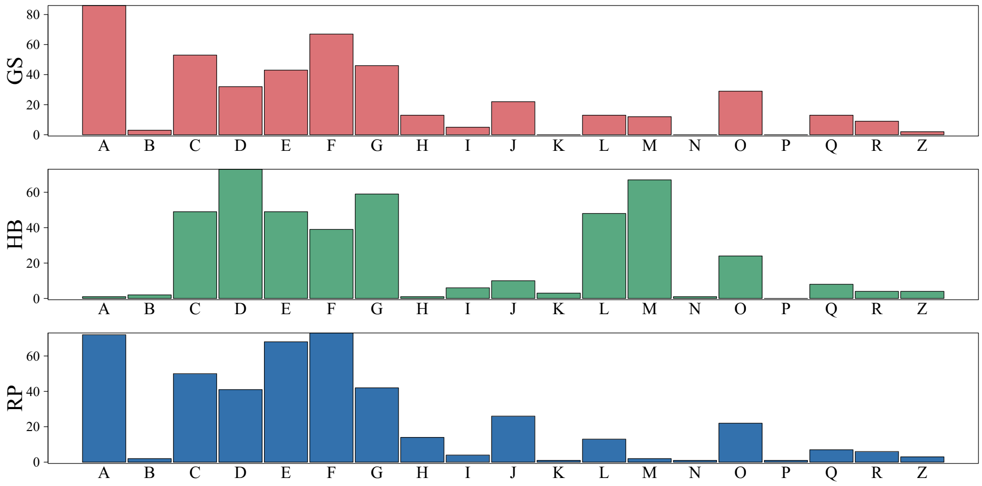

Figure 2.12: JEL codes and ranking scores of GS (upper/red), HB (middle/-green) and RP (lower/blue) for the top 458 scientists within each ranking system for December 2015.

scientists from economic systems (P), economic history (N), as well as law and economics (K), in the top positions of the discussed ranking systems is rather uncommon.

2.4 Conclusions

In summary, the comparison of academic ranking scales reveals useful information across ranking systems. Quantile regression successfully imputes the ranking data in the Handelsblatt rankings. The proposed HB common score can be used for the prediction of HB sub-ranking based on available HB data and in an inter-dependence comparison of HB, RP and GS. We have demonstrated that different correlation structures between the underlying sub-rankings exist.

The empirical results show that academic ranking variation is sensitive to age. The rank of younger and advanced-aged scientists increases more significantly than that of middle-aged researchers. Individuals from mathematical and quantitative methods occupy the leading positions across all three of the

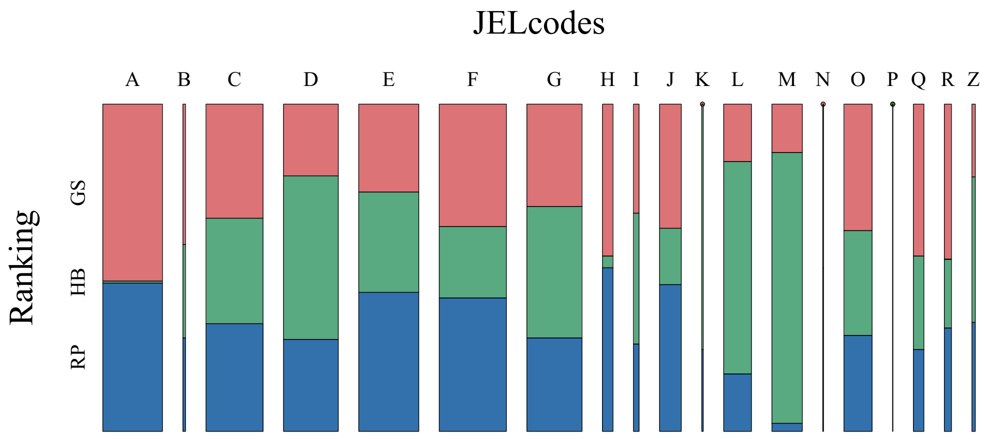

Figure 2.13: Mosaic plot of JEL codes and ranking scores of GS (upper), HB (middle) and RP (lower) for 458 scientists within each ranking system for December 2015. The width of the columns represents the number of individuals within each research area and dots represent zero.

discussed ranking systems. Individuals from microeconomics, international economics and general economics and teaching present the dominant share within HB, RP and GS, respectively. Finally, the proposed framework successfully completes research profiles of scientists.

3 Research Groups: How to Measure Performance of a Collaborative Research Center

This paper discusses Collaborative Research Centers (CRC) – long-term university-based research institutions funded by the German Research Foundation (DFG 2018). Evaluating publicly financed research results improves transparency, helps in reflection and self-assessment, and provides information for strategic decision making. Periodic monitoring of resource use and interim results allows CRC management to keep the finger on the pulse and to react to unfavourable phenomena promptly or to develop options for improvement; thereby, supporting success of the CRC.

There are numerous studies that concentrate on the evaluation of university research or research institutions in general (Pastor et al. 2015, Van den Berghe et al. 1998). Lee (2010) and Bolli and Somogyi (2011) discuss performance measurements for departments and research units. Jansen et al. (2007) and Carayol and Matt (2004) further investigate performance indicators for research groups. However, a CRC differs from common research units or institutions, because of its interdisciplinary background. The performance indicators used for the evaluations of a CRC should be designed specifically for its needs and purposes in order to reflect the behavior of involved research fields and other underlying characteristics.

In this chapter we focus on a selection of performance indicators for intermediate and final evaluations suitable for broad applicability within CRCs and identifying a relationship between productivity and resource use of CRCs that may have implications for funding policy. The goals of this paper include: (i) Selecting performance indicators suitable for a CRC; (ii) Visualizing goals vs.

results, societal impact and the interdisciplinarity structure of research results of a CRC; (iii) Analysis of a dependence structure between financial inputs and research outputs of a CRC and development of research productivity over time.

To achieve these objectives, we use the twelve years (2006 – 2012) of data from a Collaborative Research Center 649 "Economic Risk" (CRC 649) on 32 sub-projects (SPs). For each SP we observe the yearly staff costs, travel costs, number of discussion papers (DPs) and the total number of citations per year. The life span of each SP varies, which results in an unbalanced panel.

Schröder et al. (2014) indicate that the proposal for funding determines objectives for the research activity. To examine the correspondence between objectives and research results of the CRC, we carry out a semantic analysis of proposals and abstracts from published DPs. As a result, we find that both use 50% of the same words.

Apart from research activity, a CRC has an impact on society through public events, transfer of knowledge or promotion of young researchers. For instance, young researchers usually perform specific theoretical or practical research that is also used for their Ph.D. thesis. Collecting data on their further career helps to better understand this impact. With the help of a mosaic plot, we visualize three important dimensions of young researchers careers after receiving their Ph.D. within the CRC: gender, location and the area of work. For example, we show that nearly 70% of individuals obtain a job in academia.

Through a network analysis, we illustrate the interdisciplinarity structure of the research results and find out that most DPs were published in the fields of mathematical and quantitative methods, followed by financial economics, macroeconomics and monetary economics.

To study the relationship between research outcomes and funding for the CRC, we regress the number of DPs on staff and travel costs using SP-level data. With the help of year dummy variables added to the model, we show how the pattern of the SPs' productivity changed from 2006 to 2016 after controlling for staff and travel costs. Since the level of spending from the previous year and the preceding number of DPs may influence the current number of DPs, we additionally control for the lagged variables. The productivity of each SP may differ due to some heterogeneity or individual effects, such as the skills of a principal investigator (PI), average abilities or skills of researchers employed at the SP, or the specific behavior of a research field. For instance, working on a publication with one vs. more co-authors, writing in English vs. other languages, or publishing in books vs. articles may affect the research outcomes. Therefore, we allow for the possibility of individual SP effects. Considering the data structure, we apply a time fixed effects panel data (FE) model. Since the number of DPs is a count variable, we also apply a fixed effects Poisson (FEP) model.

We show that the increase of staff costs by 100% leads to an expected increase in the number of DPs by roughly 43% (FE) or 1.62 DPs (FEP). Travel costs have a diminishing effect on the number of DPs according to estimation results of the considered models. The previous level of both staff and travel costs negatively influence the number of DPs. We depict the estimates of coefficients of the dummy variables for years and find that the development trend corresponds with the stages of a project's life cycle. For instance, the most significant declines in the number of DPs take place during the stage of theoretical and empirical research, whereas the finalization stage corresponds with the growth in the number of published DPs.

The programmed R codes are available on the web-based repository hosting service and collaboration platform GitHub.

The remainder of this chapter is structured as follows. Literature review on performance indicators is presented in Section 4.2. Section 4.3 describes the data and provides some preliminary descriptive analyses. Section 4.5 introduces the methodology and shows empirical results. Finally, Section 3.4 summarizes the results.

3.1 Selection of Performance Indicators

The combination of a peer-reviewed process and quantitative indicators is common practice in research performance assessment. The German Council of Science and Humanities (WR, germ. - Wissenschaftsrat) suggests evaluating the research institutions within three dimensions (research, promoting young researchers and knowledge transfer), which contain nine research performance criteria (WR 2004). We select five criteria relevant to a CRC and provide a literature review on suitable indicators that may reflect the performance of the CRC.

1. *Research quality* shows originality and novelty of research outputs, trustworthiness of methodology, impact and relevance for further research.

Indicator	Definition	Literature
Relative reception success		
C_{Pub}	Relation of total number of citations (NC_{Pub}) to the total number of publications (N_{Pub})	WR (2012), Diem and Wolter (2013), Donner and Aman (2015)
C_{Pub}/FC_m	Citations per publication in relation to the citation's average of the field	WR (2012), Abramo and D'Angelo (2011), Moed et al. (2011), Van den Berghe et al. (1998)
C_{Pub}/JC_m	Citations per publication in relation to the citation's average of the journal	Moed (2010), WR (2012)

Table 3.1: Research quality.

2. *Effectiveness* reflects the contribution of all SPs to the development of expertise in the research field within the CRC and beyond.

Indicator	Definition	Literature
Research activity		
N_{Costs}	Total amount of the third party expenses (TPE)	WR (2012), Schmoch and Schubert (2009)
N_{Staff}	Total number of staff financed from TPF	Carayol and Matt (2004), WR (2012)
RA_{unit}	Research activity of unit (SP) – multiplication of the total number of publications and the total number of citations of a unit with regard to the institutions-wide number of citations for the analyzed period ($RA_{SP}=N_{Pub_{SP}}*C_{Pub_{SP}}/C_{Pub_{CRC}}$)	Pastor et al. (2015)
Research productivity		
N_{Pub}	Total number of publications	WR (2012), Abramo and D'Angelo (2011), Diem and Wolter (2013), Moed et al. (2011), Hornbostel (1991)
NC_{Pub}	Total number of citations	WR (2012)

FN_{Pub}	Fractional productivity – total number of contributions to publications, where each contribution is a publication divided by the number of co-authors	Abramo et al. (2009), Abramo and D'Angelo (2011)
ScS_{Pub}	Scientific strength – weighted sum of publications authored by each person, where the weights for each publication is the number of citations per publication in relation to the citation's average of the field (C_{Pub}/FC_m)	Abramo and D'Angelo (2011), Abramo et al. (2009)
h	h-index	Hirsch (2005), Bornmann (2013)
Visibility of CRC		
$AbsC_{Pub}$	Absolute citation count in the light of maximum citation count of a single publication ($C_{Pub_{max}}$) and the number of non-cited publications (N_{ncPub})	WR (2012)
Reputation		
	List of scientific prizes and awards	Zheng and Liu (2015), WR (2012)
Professional activity		WR (2012)
	Editorships Review activities Editorial board memberships Academic functions Academic memberships Organized conferences and workshops	

Table 3.2: Effectiveness.

3. The *efficiency* criterion describes a quantity of research outputs in relation to a specific input, i.e. total costs, staff expenditures, number of staff, etc.

Indicator	Definition	Literature
N_{Pub}/N_{Staff}	Relation of the number of publications (N_{Pub}) to the number of research staff (N_{Staff})	Pastor and Serrano (2016), WR (2012), Abramo and D'Angelo (2011)
NC_{Pub}/N_{Staff}	Relation of the number of citations of publications (N_{Pub}) to the number of research staff (N_{Staff})	WR (2012), Lee (2000)
N_{Costs}/N_{Staff}	Relation of the TPE to the total number of research staff (N_{Staff})	WR (2012), Pastor and Serrano (2016), Barra and Zotti (2016)

Table 3.3: Efficiency.

4. *Research enabling* relates to scientific activities that facilitate and support the research of young researchers.

Indicator	Definition	Literature
Promotion of young researchers		
N_{YR}	Total number of positions for young researchers	WR (2012)
$N_{Ph.D.}$	Total number of defended Ph.D.	WR (2012), Diem and Wolter (2013), Grözinger and Leusing (2006), Schmoch and Schubert (2009)
$D_{Ph.D.}$	Average duration of Ph.D. study	WR (2004)
$N_{Pub_{Ph.D.}}$	Total number of publications by young researchers	WR (2004)
	List of awards and prizes of young researchers	WR (2012)
	List of calls and appointments for young researchers	WR (2012)

Table 3.4: Research Enabling / Promotion of young researchers.

5. *Knowledge transfer* defines the transfer of research results and products or distribution of knowledge.

Indicator	Definition	Literature
N_{Pat}	Number of patents	WR (2011), Carayol and Matt (2004)
	List of Transfer projects	
	List of activities in public relations	WR (2012)
	List of research products and teaching materials	WR (2012)

Table 3.5: Knowledge Transfer.

3.2 Data

To provide empirical evidence, we use data from a Collaborative Research Center 649 "Economic Risk" (CRC) that was launched in 2005 for a four year term and extended twice, for a total life span of twelve years. As an interdisciplinary research center, it combined economics, mathematics and statistics and pursued research in three principal areas: financial markets and risk assessment, individual and contractual answers to risk, macroeconomic risks. For more information, we refer to the website of the CRC (CRC 649 2016).

Since the 32 sub-projects (SPs) of the CRC had different life periods, the dataset does not have the observations for all years that indicates an unbalanced panel, see Figure 3.1. The main reason for the panel being unbalanced is the attrition of SPs, as a result of research project's termination or the leave of principal investigators to other universities, and the establishment of new research projects during the prolongation phases.

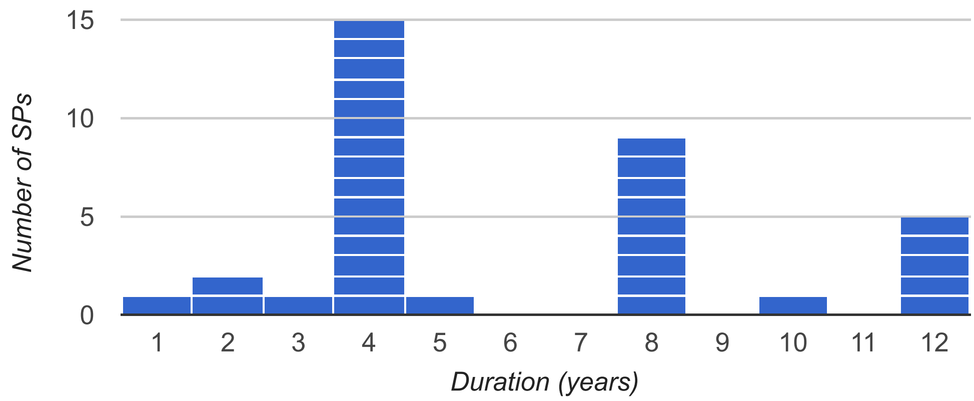

Figure 3.1: Distribution of SP life span in years.

We use data from annual financial reports, internal publications and DP's databases and CRC's newsletter. Additional insight is gathered from the texts of one proposal for a launch and two proposals for a prolongation of the CRC 649 (2005–2008, 2009–2012, 2013–2016) which were submitted to the DFG. On the one hand, one can see such proposals as goals that a CRC sets for each period. On the other hand, the published DPs encompass the achieved results of the research activity. We undertake a semantic analysis on both informational sources, i.e. 61 summaries of SPs from three proposals and abstracts of 771 DPs. The two word clouds of the top 75 keywords are illustrated in Figure 3.2. We find that both use 50% of the same words. The different size of the same words, for instance the word "risk", indicates that the number of times the word is mentioned in the proposals and abstracts differs.

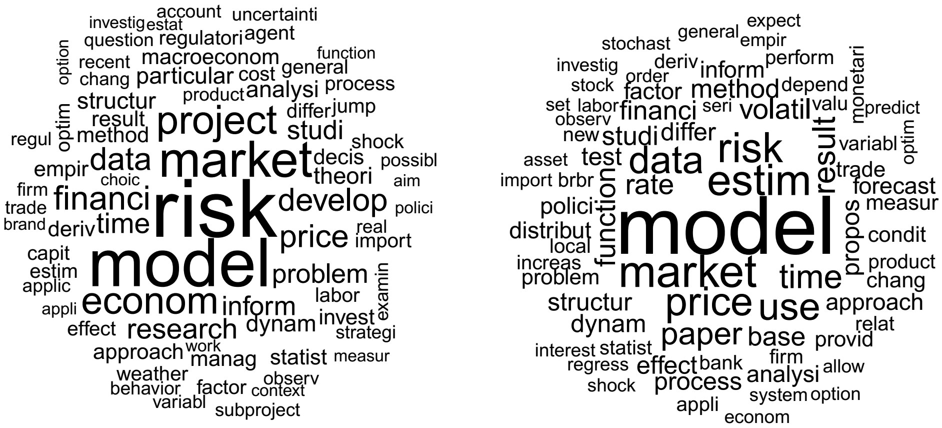

Figure 3.2: Semantic analysis of goals (left; 61 summaries from SP of three proposals for the CRC) vs. results (right; 771 abstracts from DP).

One of the primary goals of a CRC is the high-quality instruction, supervision and support of young researchers. The common result of this process is a Ph.D. defense. Collecting data on the further career of the young researchers helps to better understand the impact on society. For instance, one may wonder how many females that worked and defended their Ph.D. thesis in CRC are afterward working in academia in Germany? To visualize such data we use a mosaic plot in Figure 3.3.

The vertical axis splits the individuals according to their gender. The data are further divided into two groups on the upper horizontal axis according to the location of the job. The lower horizontal axis shows how many people received a contract in academia or other fields. The width and height of each segment represent the number of observations within each group. Consider the 65 members of the CRC that received their Ph.D. from 2005 to 2016. There are 11 female researchers that received jobs in academia in Germany and 6 in other countries. For males that stayed in academia, the number is 21 for Germany and 7 for other countries. This means that around 70% of young researchers employed by the CRC have found a job in a scholarly institution.

In order to understand if the intended interdisciplinarity occurred, we analyze DPs that serve as an outcome of the CRC research activity. Almost each DP has codes indicating subject fields according to the Journal of Economic

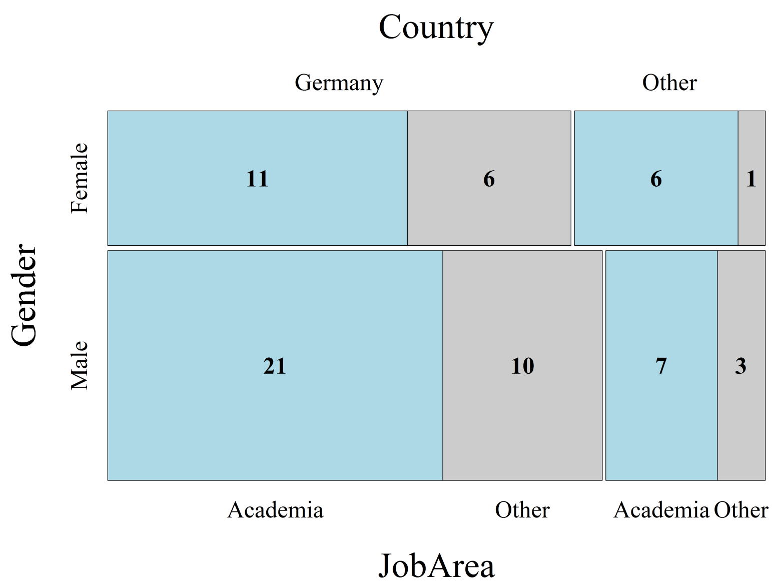

Figure 3.3: Mosaic plot of job type, location and gender of 65 CRC members who received their Ph.D. between 2005 and 2016 (as of Dec 2016).

Literature (JEL) classification in the economic sciences (see JEL 2018).

We show the network of collaborating disciplines in Figure 3.4. The small gold circles introduce the DPs, whereas the nodes leading to the bigger blue circles indicate the JEL code of the corresponding research area. The size of each blue circle reflects the relative number of references to DPs. The explanation of JEL codes is given in Table 2.4. For instance, most of the DPs were published in the C area, i.e. mathematical and quantitative methods. They are followed by G, financial economics, and E, macroeconomics and monetary economics.

3.3 Analysis of Research Productivity

The observed time series across the same SPs indicate the longitudinal or panel structure of the data. To investigate the relationship between the input and the output variables, we use the methods designed for panels.

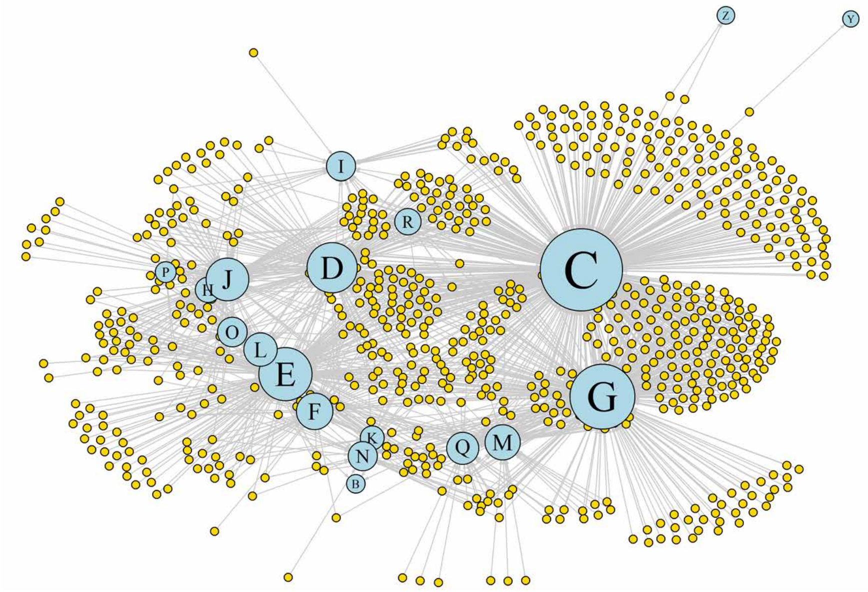

Figure 3.4: Network of 760 discussion papers (yellow) and 20 JEL codes (blue) published from 2005 to 2016.

3.3.1 Methodology

The basic framework for the panel data analysis shows the model (Wooldridge 2002):

$$y_i = \beta X_i + u_i, \quad i = 1, \dots, K, \tag{3.1}$$

where $y_i = (y_{i1}, \dots, y_{iT})^\top$ is a $(1 \times T)$ vector of observations for $t = 1, 2, \dots, T$, $X_i = (x_{i1}^\top, \dots, x_{iT}^\top)^\top$ is a $(K \times T)$ matrix of observations, β is a $(K \times 1)$ vector of coefficients and u_i is a $(1 \times T)$ vector of unobservables.

The unobserved SP effect may contain such factors as publishing behavior in a research field, average researchers' abilities or skills of principal investigators of SPs that should be roughly constant over time.

We allow for arbitrary correlation between the unobserved SP heterogeneity or fixed effects c_i and the observed explanatory variables x_{it} and, therefore, use the fixed effects model for each i (Wooldridge 2016):

$$y_{it} = \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + c_i + u_{it}, \quad t = 1, 2, \ldots, T, \quad i = 1, 2, \ldots, K, \tag{3.2}$$

where y_{it} includes dependent variables and x_{it} independent variables for individual i at time t, $\beta_1, \ldots, \beta_k$ are the unknown coefficients, c_i is individual effect or individual heterogeneity and u_{it} are idiosyncratic errors that change across individuals i and time t.

The fixed effects estimator (or the within estimator) is obtained as the pooled OLS estimator on the time-demeaned variables. The strict exogeneity assumption on explanatory variables, $\mathrm{E}(u_{it}| \mathbf{X}_i, c_i) = 0$, provides that the fixed effects estimator is unbiased (Wooldridge 2016). As the number of SPs (clusters) is large, statistical inference after OLS should be based on cluster-robust standard errors to account for heteroscedasticity and within-panel serial correlation (Cameron and Miller 2015).

Next, we are interested in the pattern of SPs' productivity, i.e. number of produced discussion papers, in different time periods. For this purpose we use time fixed effects that change over time but are constant across SPs. We include the dummy variables for $T - 1$ years to avoid the multicollinearity. Usually the first year is selected as a base year. The time fixed effects model (FE) is (Stock and Watson 2003):

$$y_{it} = \beta_1 x_{it1} + \ldots + \beta_k x_{itk} + \delta_1 + \delta_2 D_2 + \ldots + \delta_T D_T + c_i + u_{it}, \tag{3.3}$$

where $D_2, \ldots, D_T$ are time effects and $\delta_1, \ldots, \delta_T$ are the parameters to estimate.

When the dependent variable involves count data, it has a Poisson distribution instead of a normal distribution. Hausman et al. (1984) introduce a fixed effects Poisson model (FEP) as:

$$\mathrm{E}(y_{it}|x_i, a_i) = a_i \mu(x_{it}, \beta_0), \quad t = 1, 2, \ldots, T, \tag{3.4}$$

where β_0 is a $(1 \times K)$ vector of unknown parameters to be estimated and μ is the conditional mean. Wooldridge (1999) further derives a consistent estimator for FEP using a quasi-conditional maximum likelihood estimator (QCMLE).

3.3.2 Empirical Results

Before presenting the estimates, we explain some specifications of the model. Since the yearly staff and travel costs are in nominal Euros, a slight increase may happen due to inflation. One possibility to deal with this is an adjustment using a Consumer Price Index (CPI). Another way to track the effect of real spendings is the use of a logarithmic form. The interpretation of the estimation results is then done using the level-log model.

Table 3.6 presents the results of FE (1) and (2), and FEP (3) and (4) models for the number of DP as a dependent variable. The parameters of interest are staff costs $\beta_{logStaffCosts}$, travel costs $\beta_{logTravelCosts}$ and year-specific influence δ_{year}. We also include lagged variables into the models (2) and (4), since the current number of research outputs may be affected by the previous number of publication and invested funds in economic sciences and mathematics. The models (2) and (4) encompass the number of DPs $\beta_{nDP_{t-1}}$, staff costs $\beta_{logStaffCosts}$ and travel costs $\beta_{logTravelCosts}$ in the time $t-1$. The intercept *const* is the average of individual effects c_i across all SPs that is reported by Stata. We use cluster-robust standard errors to account for heteroscedasticity. The significance level of all estimates decreases as a result of standard error adjustment (Wooldridge 2016).

In (2) and (4) two years were omitted because of collinearity. In (3) five observations were dropped out of the analysis because there was only one observation per group. Performing analysis on unbalanced data slightly increases the estimated effects of considered variables, but the general idea remains unchanged (Wooldridge 2016).

In the model (1) we see the positive, significant effect of staff costs on the number of DPs. 1.38/100 is the unit change in nDP when staff expenses increase by 1%. In other words, a 100% increase in staff costs leads to an increase in the number of DPs by 1.38. Similarly, the model (2) shows that a

Dependent variable: nDP	FE model		FEP model	
	(1)	(2)	(3)	(4)
$\beta_{logStaffCosts}$	1.38^{**}	1.62^{*}	0.47^{***}	0.43^{**}
	(0.61)	(0.88)	(0.12)	(0.19)
$\beta_{logTravelCosts}$	-0.94^{*}	-0.34	-0.22^{**}	-0.04
	(0.55)	(0.47)	(0.10)	(0.09)
δ_{2006}	1.61	1.92	0.25	0
	(1.36)	(1.61)	(0.26)	(*omit.*)
δ_{2007}	-1.20	-2.55	-0.30	-0.98^{***}
	(1.38)	(2.46)	(0.31)	(0.25)
δ_{2008}	-0.95	-2.03	-0.23	-0.97^{***}
	(1.30)	(2.10)	(0.32)	(0.36)
δ_{2009}	-2.05^{*}	-3.16	-0.54^{*}	-1.20^{***}
	(1.13)	(1.98)	(0.33)	(0.23)
δ_{2010}	-1.93^{*}	-2.13	-0.51^{*}	-1.03^{***}
	(1.14)	(2.68)	(0.30)	(0.31)
δ_{2011}	1.10	0	0.33^{*}	0
	(0.70)	(*omit.*)	(0.20)	(*omit.*)
δ_{2012}	-2.79^{*}	-3.60^{*}	-0.71^{**}	-1.90^{***}
	(1.46)	(1.78)	(0.34)	(0.20)
δ_{2013}	-2.98^{**}	-3.18	-0.80^{**}	-1.32^{***}
	(1.30)	(2.52)	(0.32)	(0.41)
δ_{2014}	-1.36	-1.73	-0.44	-0.99^{***}
	(0.95)	(1.61)	(0.27)	(0.37)
δ_{2015}	-2.55^{**}	-1.90	-0.74^{**}	-1.02^{***}
	(1.17)	(1.77)	(0.33)	(0.31)
δ_{2016}	-0.30	0	-0.31	-0.69^{*}
	(1.79)	(*omit.*)	(0.36)	(0.41)
const	-2.37	0.05		
	(5.29)	(10.09)		
$\beta_{nDP_{t-1}}$		0.02		-0.01^{*}
		(0.16)		(0.03)
$\beta_{logStaffCosts_{t-1}}$		-0.66		-0.25
		(0.59)		(0.23)
$\beta_{logTravelCosts_{t-1}}$		-0.21		-0.02
		(0.58)		(0.13)
R^2	0.20	0.21		
AIC	706	437	463	253
BIC	742	469	501	258

Table 3.6: Estimation results for time fixed effects (within) regression (models (1) and (2)) and fixed effects Poisson regression (models (3) and (4)) with number of DP (nDP) as the dependent variable and with robust standard errors adjusted for clusters in SP.

100% increase in staff costs increases the number of DPs by 1.62, holding other variables constant. The fit of the FE models in (1) and (2) in Table 3.6 with nDP as the dependent variable is almost the same, indicating that including lagged variables does not significantly improve the model.

The FEP estimates have a different interpretation. For instance, the coefficient on $\beta_{logStaffCosts}$ shows that a rise of staff costs by 100% leads to an increase of the number of DPs by 47% and 43% for models (3) and (4) correspondingly. The coefficients on staff costs estimates for four models in Table 3.6 are significant at 1% to 10% level. The influence of previous values of staff costs on the number of DPs is negative and insignificant.

Travel costs have a diminishing effects on the number of DPs according to estimation results of considered models. The coefficient on $\beta_{logTravelCosts}$ implies that, if we increase the travel costs by 100%, we expect the number of DP to decrease by 0.94 DP due to FE model (1). The Poisson coefficient in (3) means that an increase in $logTravelCosts$ by 10% decreases nDP by 2% (0.22×0.10).

The coefficients on the year dummy variables reveal how the average productivity of SPs changes over time. As 2005 is selected as the base year, it is not reported with a coefficient. The coefficient on δ_{2006} in model (1) shows that, on average, 1.6 DPs are attributed to the year effect of 2006 holding other factors fixed. In Poisson case (3) one suggests that the expected number of DPs in 2006 is 25% higher than on average. The coefficients on δ_{2006} and δ_{2011} indicate a positive increase in the number of DPs even without changing expenses. The omission of year dummies would lead to the attribution of this positive effects to the effects of costs change.

One can see that the year effects have a negative impact on the number of DPs in the majority of years for all models. The project's life cycle could explain this. Research projects generally have five main stages: proposal development, funding review, project start-up, performing research and finalization of the project. We map the estimates of coefficients of the models and fit the stages of life cycles in Figures 3.5 and 3.6. Proposal development and funding review take place before 2005 and are not depicted in these Figures.

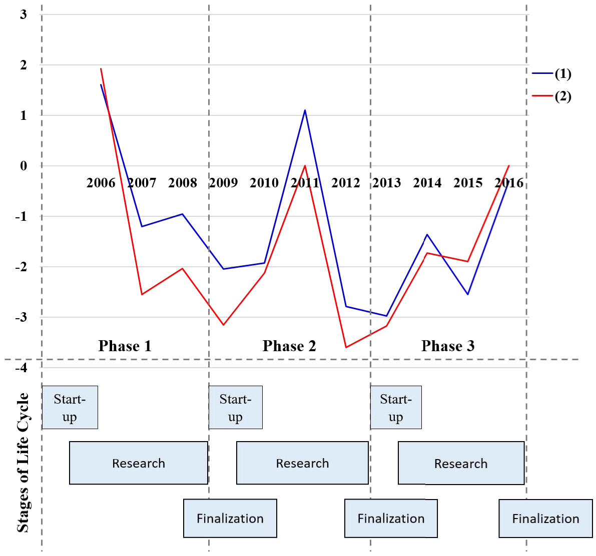

Figure 3.5: Estimates of coefficients on the year dummy variables for FE models. The lower part of the figure shows the corresponding stage of the research project life cycle.

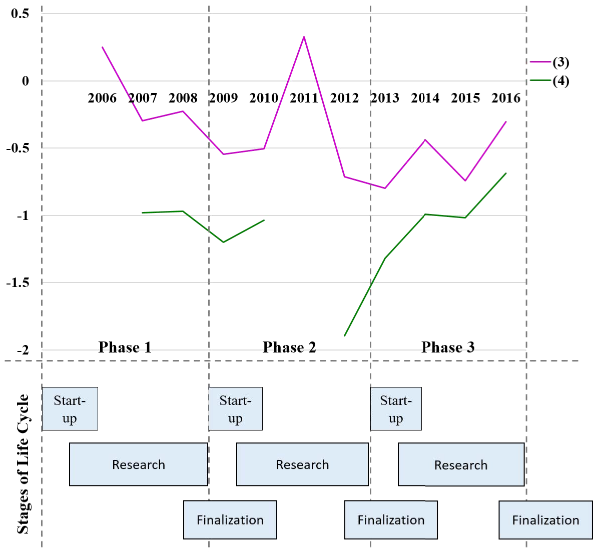

Figure 3.6: Estimates of coefficients on the year dummy variables for FEP models. The lower part of the figure shows the corresponding stage of the research project life cycle.

A highly demanding application for a CRC requires extensive preliminary research. The results of this preliminary research are published as DPs in the first year 2005, thus, creating a specific bias towards later research outputs produced during the CRC's life time. The three following increases in the number of DPs take place in the finalization stage caused by the publishing of research results at the final stage of projects. The research outputs of the last phase in 2016 show part of the positive trend. In fact, 28 DPs were published in 2017, after the CRC was officially finished and financing ended. Three major declines could be explained by a theoretical and empirical stage of the research in the middle of each project life cycle. In summary, the joint depiction of the time fixed effects and the research project's life cycle could allow a better understanding of the development of the number of DPs over time.

3.4 Conclusions

Our findings show that the performance indicators suitable for the intermediate or final evaluation of a CRC help create a better understanding of the dependence structure between research productivity and financial inputs, and provide relevant information for successful decision and policy making. Using time fixed effects panel data model and fixed effects Poisson model, we show that increasing staff costs by 100% raises the number of DPs by 1.62 or 43% according to the estimates of FE and FEP models correspondingly. Travel costs have diminishing effects on the number of DPs according to our estimation results. We analyse the change in productivity of CRC over time for reasons not captured by the other independent variables using the dummy variables for years. We depict the estimates of coefficients for years and show that the development trend could correspond to the stages of a project's life cycle. For instance, the major declines in the number of DPs take place during the stage of theoretical and empirical research, whereas the finalization stage may correspond to the growth in the number of published DPs.

4 Universities: Is Scientific Performance a Function of Funds?

This chapter contributes to a deeper understanding of the interplay between third-party funds (TPF), publications and citations. The resulting analysis reveals the significance of generally accepted beliefs in the scientific community and provides guidelines for improvement of decision making in university research management.

Many studies investigate the relationships between bibliometric indicators on an individual level (Abramo et al. 2013, Wildgaard 2016, Costas et al. 2010). However, an analysis of both bibliometric outputs and grant budget data mainly uses data at an aggregated level, i.e. institutes, faculties or departments (Bolli and Somogyi 2011, Dyckhoff et al. 2009). Few studies compare the bibliometric indicators with funds on the micro-level of research groups or university chair level (Jansen et al. 2007, Carayol and Matt 2004, Rosenbloom et al. 2015). Undertaking analysis on the level of individuals and subsequently merging outcomes into groups of interest could yield more robust and reliable results. Ebadi and Schiffauerova (2016) perform such statistical analysis of research funding of individual researchers listed in Natural Sciences and Engineering Research Council of Canada (NSERC) and their scientific outputs as recorded by Scopus. However, their study does not consider the financial support researchers receive from other funding institutions that may also affect the number of publications and citations. Mongeon et al. (2016) explore the distribution and marginal returns of research funding using data of the entire population of Québec academic researchers on funding, publications and average relative citations. However, they divide all researchers into three broad disciplines and perform analysis using the average and median of groups of 50 researchers. The discipline-specific characteristics are averaged within three research areas.

A distinctive feature of our study is the analysis of individual-level data from a German university, which belongs to the top 10 universities in Germany in terms of external funds acquisition (DFG 2015). A sample of professorships, the complete set of their third-party expenses (TPE), publications, and citations from Scopus, is observed on a yearly basis for the period 2001 to 2015. Additionally, we include a variable academic age (number of years after Ph.D. degree). This information enables the analysis on a fine level of granularity and provides the possibility to account for time-delayed effects.

Some researchers collaborate with their colleagues from other faculties and subsequently fields. As a result, their research outputs may reveal an interdisciplinary character that yields a heterogeneity. The analysis on individual data level sheds light on the heterogeneity of actual research outputs. To display the cooperation structure between faculties, we suggest using a chord diagram (a technique commonly used in genetic engineering for genome data) and the information on co-authorship of publications. The Sankey plot visualizes the resulting research interdisciplinarity.

Decision and policy making in research management must take into account the research fields' heterogeneity. Given thoughts about the feedback and interdependency, we employ a panel vector autoregressive model with exogenous variable (PVARX) (Canova and Ciccarelli 2013, Cavallari and D'Addona 2014). Aiming to underline the existing inter-faculty heterogeneity, we estimate the PVARX model for each faculty. The resulting impulse response functions (IRF) help to understand the relation between variables in a VAR context and clarify how a change in one variable affects another variable. For example, one may be wondering to what extent the number of publications will change, if the TPE increase by 1%. Since the analysis of such original innovations is rarely the case in work with real data (Tsay 2014), we proceed with orthogonalized innovations received using Cholesky decomposition of the white noise covariance matrix. Finally, with the help of a forecast error variance decomposition (FEVD), we demonstrate a percentage of the variance of the prediction error explained by a shock at a four-year time horizon.

Our findings inform the university research management about the interrelationship between research performance indicators for each faculty and provide a range of possible explanations for the revealed patterns across scientific areas.

We quantify the influence of TPE, publications and citations on each other; the reaction of the system to exogenous impulses; and the amount of variance explained by considered variables. This perspective suggests the possibility to leverage the key resources according to the fields' needs and desired outputs. We summarize the results for social sciences and humanities, life sciences and mathematical and natural sciences. We also address the possible implications for policy and decision making and propose recommendations for university research management.

4.1 Literature Review

4.1.1 Third-Party Funds

TPF are financial input to the university or other institution from external sources on top of the regular university funds (Hornbostel 2001). An interesting variable is the amount of third-party funds (TPF) that were spent – third-party expenses (TPE) – because the unused part of TPF generally must be returned to the funding agencies on a yearly basis.

The academic community debates the use of TPF or TPE for a research performance evaluation. On the one hand, TPF is often accepted as an indicator of research performance, since competent experts from the corresponding subject fields carry out the peer review process before the allocation of TPF (WR 2011). However, this does not apply to all TPF. Apart from scientific funding organizations, universities receive a high amount of TPF from industry with a simplified selection process.

On the other hand, Sousa (2008) and Laudel (2005) caution against using TPE as an indicator of academic excellence. TPE measures only limited, or not at all, the quality of research or knowledge process. In contrast, bibliometric indicators and the results of the peer review process should be more appropriate for this purpose. Gerhards (2013) criticizes that in Germany the role of TPF is overemphasized in comparison to other countries. For instance, the assessment of the research quality in the United Kingdom via the Research Excellence Framework (REF) and in the United States via the National Research Council

(NRC) Ranking uses mainly publications and citations to measure the research performance of universities (REF 2011, NRC 2010). Gerhards (2013) points out that one should use TPF not as an output but as an input variable, which enables the research process and, as a result, publications, patents, inventions, etc. He further concludes that TPF measures are not suitable as an indicator of research quality, unless the correlation between TPF and research results is strong enough. Therefore, it is important to determine and understand in which research areas high TPE can be associated with high research outputs and acknowledgement among the scientific community.

Lariviere et al. (2010) emphasizes the various TPF demand in different research areas. In other words, the TPF varies significantly across research fields. For instance, natural and engineering sciences generally need expensive equipment, that is often financed through TPF, in order to start the research activity (Hornbostel 2001). At the same time, humanities and social sciences require mainly access to literature usually provided by the institution and research staff that could be financed through TPF. As concluded by Jansen et al. (2007), a significant difference emerges due to the field-specific practice of raising TPF. Thus, the absolute amount of TPE could only indicate the productivity in each field.

4.1.2 Publications and Citations

Publications and citations generally act as an indicator of research productivity and resonance of research outlets. The majority of bibliometric studies focus on articles and literature reviews. One may exclude other document types because of the difficulties in comparison (Waltmann 2016).

In the same way as for TPE, the publishing and citing behavior differs across fields. For instance, social sciences and humanities tend to publish in monographs, books or regional and national journals. Law sciences often publish in the national language. Conference papers are the basic platform for the introduction of research results for computer scientists, whereas in natural sciences and economics articles are standard. Because of the field specifics, publications in high-energy physics and biomedical sciences can count hundreds of co-authors. Therefore, researchers in these areas correspondingly have a significantly higher number of publications and citations. A more detailed overview of the publishing practice in various research areas is provided in Hornbostel and Torger (2015). Similarly, the citation behavior varies across fields. Hicks et al. (2015) and Bartol et al. (2014) introduce the significant difference in the number of citations in particular subject areas and point out the necessity to normalize.

Further, the language has direct influence on the number of citations of publication and as a result on the international visibility of research (Gerhards 2013). The question whether to include non-English publications into the corpus is arguable. Although some studies insist on including only English publications when comparing research institutions (van Leeuwen et al. 2001, van Raan et al. 2011), such an approach will penalize the scientific fields with non-English publishing behavior.

An key factor influencing the outcome of analyses is the publication count. When analyzing a publication written by two co-authors, for example, one should decide which counting method to choose: full (assigning weight 1 to each author) or fractional (weight 0.5). This choice usually depends on the objective of a particular study. Moed (2005) explains the difference between full and fractional counting as the difference between participation and contribution. Waltmann (2015) provides a comprehensive literature review on the choice of counting method.

The time delay between when the research work is published and when it begins receiving citations is called the citation window (Glänzel and Schubert 2003). The size of the citation window influences the number of publications and citations that subsequent citation analysis will use. A large size of the citation window leads to the exclusion of more recent publications from the analysis, as they do not have enough time (equal to the length of the citation window) to collect the necessary citations (Waltmann 2015). However, when the citation window is too small – for instance, one or two years – the mapping of the citations' impact can be incomplete. Setting the value of the citation window is important because it provides similar conditions for comparison of publications of different age. For instance, an article published ten years ago can collect many more citations than the one published five years ago ceteris paribus. Nevertheless, Abramo and D'Angelo (2011) study the differences between scientific areas and conclude that the long-term and short-term citation counts correlate strongly. In the case of a comparative analysis, the length of the citation window equal to three years can be used (Abramo et al. 2011). Therefore, the purpose of research should help with the choice of citation window: large window size for more accuracy or small size for stressing of timeliness (Wang 2013).

4.2 Research Model

The main TPF objective is the support of research. It is natural to assume that the scientific outcome is presented to the scientific community through publication channels. Previous studies demonstrate that the research funding has a positive impact on the knowledge production and publication output (Jacob and Lefgren 2011, Boyack and Borner 2003, Payne and Siow 2003, Rosenbloom et al. 2015, Bolli and Somogyi 2011, Carayol and Matt 2004). Using both bibliometric and regression analyses Ebadi and Schiffauerova (2015) and Ebadi and Schiffauerova (2016) confirm a strong relation between allocated funds and the productivity of researchers. According to the results of McAllister and Wagner (1981), this tendency is true for various fields of science. Furthermore, Beaudry and Allaoui (2012) identify a J-shaped curve explaining the significant positive effect of public funding on the publication rate. In other words, researchers with more funding produce even more publications. Summing up the literature, we propose the following hypothesis:

> *H1: Researchers with more funding have higher publication productivity, i.e. there is a positive impact of the past third-party funds on the current number of publications.*

The allocation of TPF to researchers, as a result of a highly competitive peer-review process, is based inter alia on the prior research work. Beaudry and Allaoui (2012), Nag et al. (2013) and Rosenbloom et al. (2015) show that past scientific productivity positively affects the likeliness of obtaining grants. A higher number of publications may result in a higher amount of acquired funding. Laudel (2005) explains this may be due to the fact that researchers applying for external funds must display some previous published research.

This suggests:

> *H2: The past productivity of researchers influences the likeliness of obtaining external funding, i.e. there is a positive effect of the number of publications in the past on the current amount of third-party funds.*

Different from the strong positive influence on the quantity, allocated funds exhibit only partly a related effect on the quality of research outputs. Mongeon et al. (2016) report that the increase in funding leads to an increase in scientific impact. In other words, citations – up to a certain level – are followed by a rapid decrease of marginal returns. They further explain that the reason may be different allocations of time, e.g. writing funding proposals or performing administrative tasks. The results obtained by Payne and Siow (2003) confirm a low and a negative relation between research funding and citations per article, suggesting that the growth in expenditures yields in a higher quantity, but not necessarily quality, of publications. This leads to:

> *H3: The academic funding of researchers influences the number of citations accumulated by their publications, i.e. there is a relationship between the amount of third-party funds and the number of citations.*

Apart from indicators of scientific performance, the age of researchers may have a positive or negative influence on the other factors. As a result of analysis of researchers' data from different scientific fields, Cole (1979) shows that there is a minor curvilinear relation between age and indicators of research performance. He concludes that this influence is, however, low. Beaudry and Allaoui (2012) provide similar results. On the contrary, the analysis of Bonaccorsi and Daraio (2003) on the micro level demonstrates that the scientific productivity decreases with increasing age of scientists. Levin and Stephan (1991) provide evidence that, on average, there is a negative relation between age and productivity of researchers. Further, Abramo et al. (2016) show this is true regardless the research area. Whereas Kyvik (1990) points out the differences in research fields for various age groups of scientists, indicating a greater decreasing trend in productivity for disciplines with frequent and extensive technical changes. Based on this review, we formulate two more hypotheses as:

H4: The scientific productivity of researchers changes with time, i.e. there is an effect of age on the number of publications.

H5: The amount of academic funding changes with the age of researchers, i.e., there is an impact of age on third-party funds.

The research model with the corresponding hypotheses is illustrated in Figure 4.1.

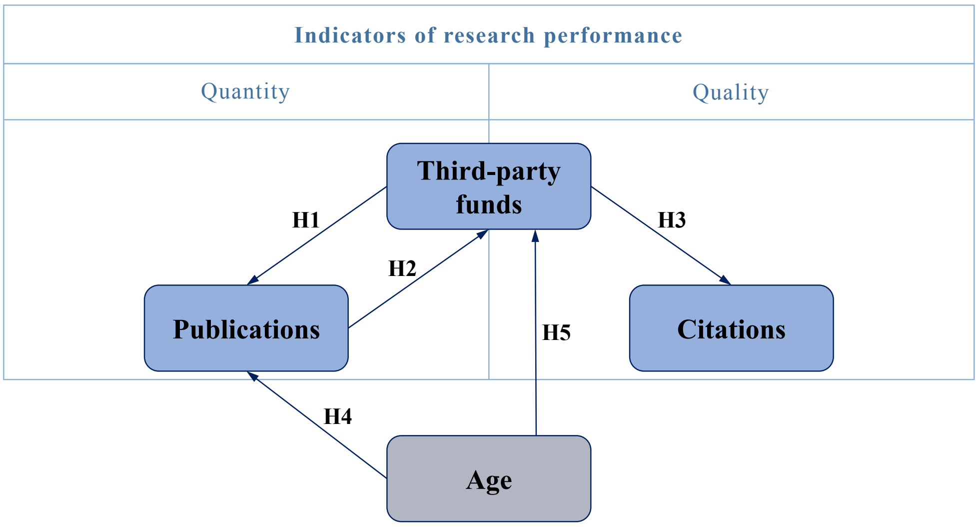

Figure 4.1: Summary of the research model and hypotheses.

4.3 Data

We obtain the individual-level data on TPE for professorships (scientific units of chair holders or lab owners) from a German university that belongs to the top–10 German universities in terms of TPF acquisition (DFG 2015) and to the top–5 German universities according to a Times Higher Education's World University Ranking 2018 (THE 2018). The data covers the period 2001 to 2015. From an organizational point of view, each professorship belongs to one of eight faculties. However, three of the faculties (Faculty of Life Sciences, Faculty of Humanities and Social Sciences, Faculty of Mathematics and Natural

Sciences) contain quite dissimilar institutes with regard to the differences in the corresponding research areas. For this reason, we split those faculties into the lower level and in the end receive 16 entities for analysis, see Table 4.1 and Figure 4.2. For simplicity we further name these entities faculties. The descriptive statistics is presented in the Appendix.

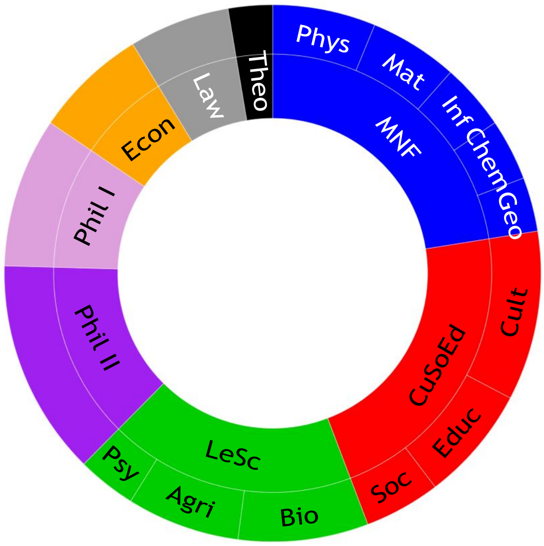

Figure 4.2: Sunburst plot for faculties and lower aggregation level. The width of segments corresponds to the number of professorships in each unit in 2015 (680 in total). The data of eight outliers are removed.

One should note that assistant professors, research assistants or other scientific members of the chair can gain their own third-party projects. Those TPE are also allocated to chair holders or lab owners, although the results of these research projects are not necessarily published under the chair's flag. The development of TPE for professorships grouped by their corresponding faculty from 2001 to 2015 is introduced in Figures 4.3 – 4.5.

Furthermore, we match each chair holder who had TPE in the period 2001 to 2015 with his or her publications and citations listed in the Scopus database. We choose the Scopus database as a source for publications and citations, as it is currently the most extensive database of academic literature and provides better coverage of publications and citations for the majority of disciplines

Original faculty		Analysed unit (faculty)	
Abbr.	Full name	Abbr.	Full name
Social sciences and Humanities			
Law	Faculty of Law	Law	
Phil1	Faculty of Arts and Humanities	Phil1	
Phil2	Faculty of Language, Literature and Humanities	Phil2	
Theo	Faculty of Theology	Theo	
Econ	Faculty of Economics and Business Administration	Econ	
CuSoEd	Faculty of Humanities and Social Sciences	Cult	Cultural History and Theory, Art and Visual History, Musicology and Media Studies, Archaeology, Asian and African Studies
		Soc	Social Sciences, Transdisciplinary Gender Studies
		Educ	Education Studies, Sports Sciences, Rehabilitation Sciences
Life sciences			
LiSc	Faculty of Life Sciences	Agri	Agriculture and Horticulture
		Bio	Biology
		Psy	Psychology
Mathematical and Natural Sciences			
MNS	Faculty of Mathematics and Natural Sciences	Chem	Chemistry
		Geo	Geography
		Inf	Computer Science
		Mat	Mathematics
		Phys	Physics

Table 4.1: Organisational structure of analysed data.

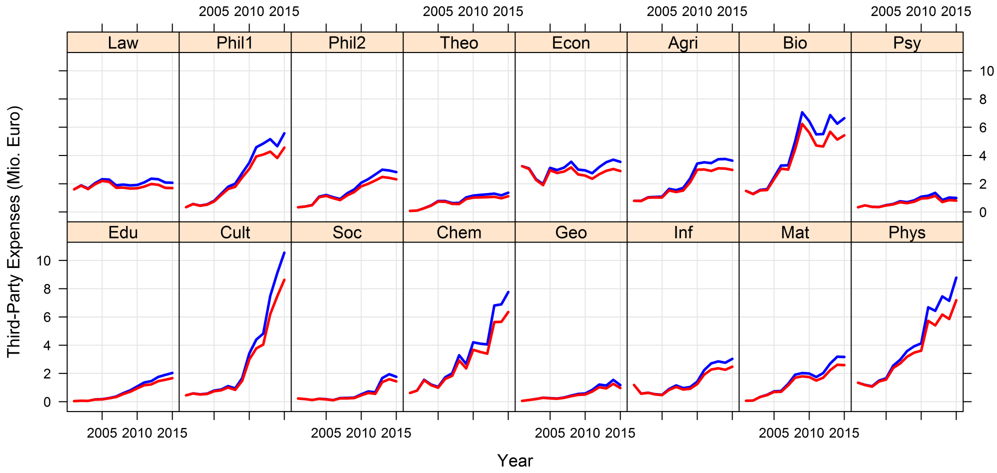

Figure 4.3: Total amount of TPE of professorships from 2001 to 2015. The data of eight outliers are removed. The nominal value (blue) and the inflation adjusted real value (red).

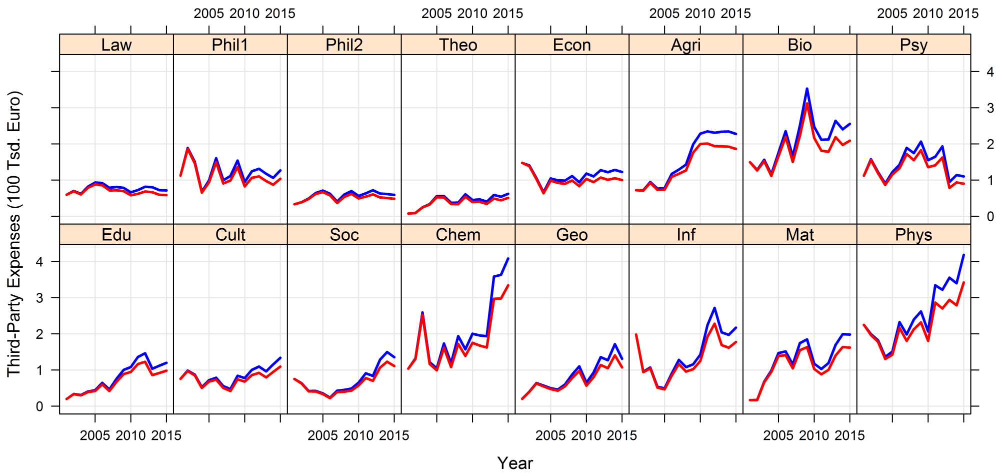

Figure 4.4: The development of nominal (blue) and inflation adjusted real (red) TPE in relation to the number of professorships with TPE within each faculty from 2001 to 2015 without eight outliers.

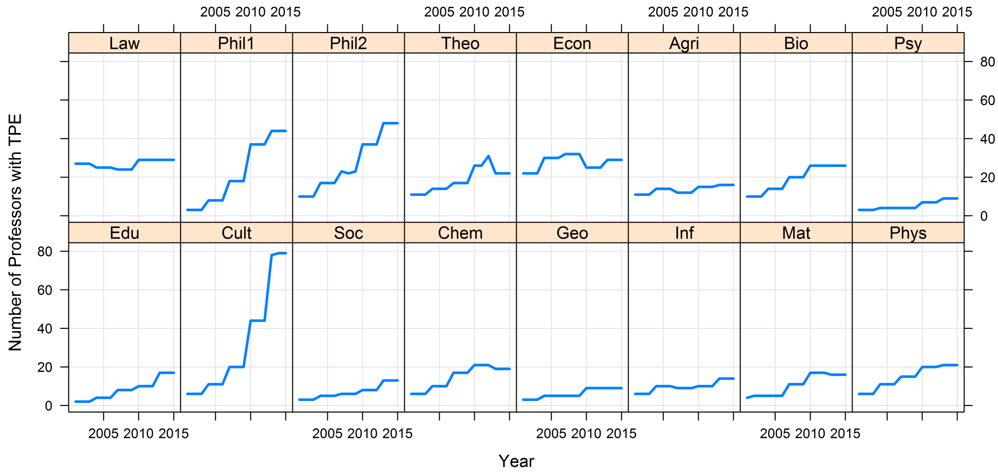

Figure 4.5: HU professors with TPE through the faculties from 2001 to 2015. The data of eight outliers are removed.

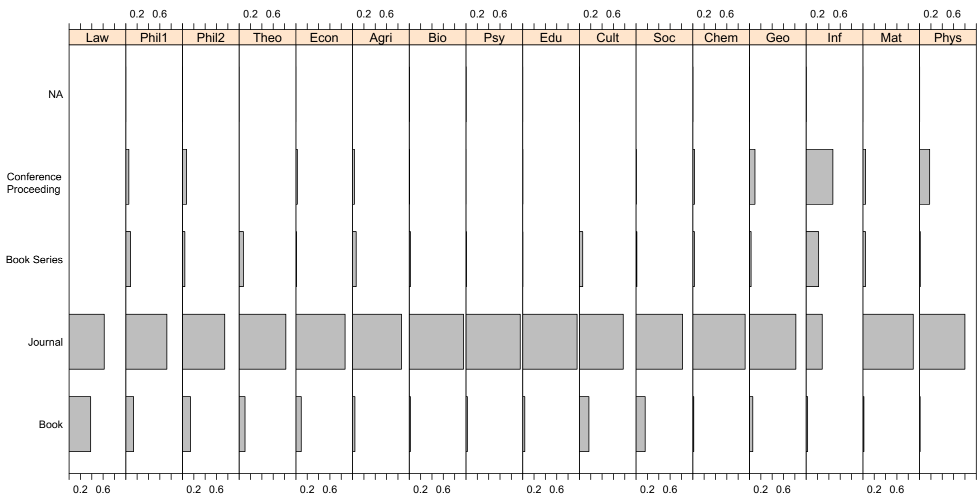

Figure 4.6: Frequency of publications of each document type published by professors grouped by faculties from 2001 to 2015. The data of eight outliers are removed.

compared to a Web of Science database by Thomson Reuters (Bergman 2012, Bartol et al. 2014, Waltmann 2015). The census date of the Scopus citation data is 31.08.2017. As a part of data preparation, we select publications of all document types for the selected corpus of authors, namely articles, conference papers, literature reviews, chapters, editorials, articles in press, errata, notes, books, letters, short surveys and conference reviews. Hereby, we capture the outlets common for different subject fields. For the most faculties, articles form the basis of the corpus (see Figure 4.6). A well-known exception is computer science with a large part of conference proceedings, whereas law, social sciences and humanities outputs are found mostly in books, book chapters and conference proceedings. As a matter of fact, law, theology and social sciences have noticeably fewer recorded Scopus publications in comparison to other faculties.

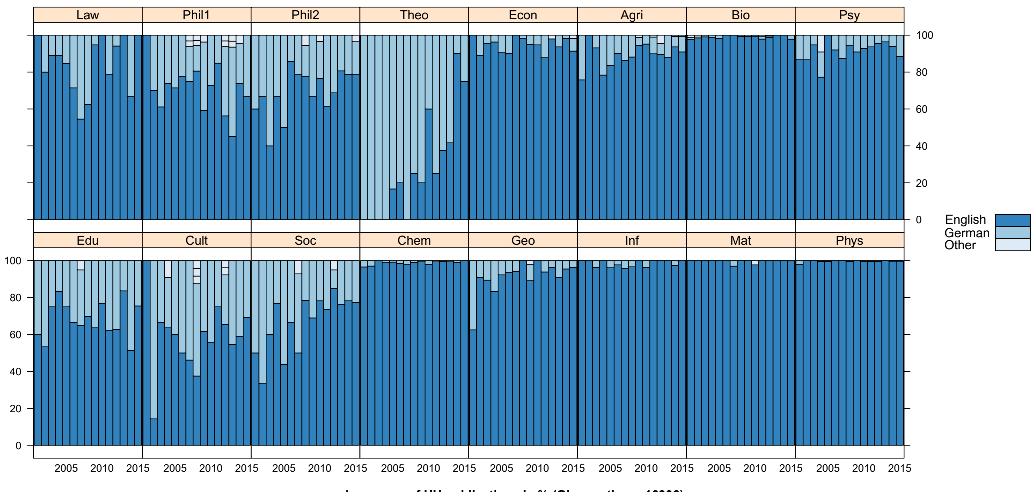

Figure 4.7: Proportion of languages (EN – dark blue, DE – blue, others – light blue) of all publications in corpus from 2001 to 2015. The data of eight outliers are removed.

As we are interested in the overall performance of faculties from the perspective of participation, we select the full counting method. We include all languages of indexed publications to avoid penalization of scientific areas with non-English publishing behavior – for instance, social sciences and humanities. An overview of the proportion of languages of all publications from 2001 to 2015 is introduced in Figure 4.7. We also remove 8 outliers (3 from Biology, 5 from Physics) that have more than 100 co-authors within a single publication, as they are likely to distort the results. The development of the number of publications and citations per person over time for faculties is illustrated in Figures 4.8 – 4.9.

A closer look at the database reveals that the average number of co-authors of publications differs among faculties, see Figure 4.11. Therefore, when the cooperation structure through co-authorship is a point of interest, the research areas where fewer co-authors are common, e.g. social sciences and humanities, give less evidence for analysis. In contrast, the fields with several and more co-authors, e.g. natural and life sciences, provide a sound basis for further investigations, see Figure 4.12.

For illustration of internal (intramural) collaboration within a university, we suggest using a chord diagram in Figure 4.13. One identifies joint publication channels among university members. The scale indicates the number of publications in the period 2001 to 2015. In the left panel we use grey lines to depict the collaboration across faculties. The corresponding connections are relatively thin. This indicates that the collaboration within faculties prevails, whereas cross-faculty collaboration is less common. After excluding the publications within faculties, the research channels between faculties are more visible in Figure 4.14. For instance, there are nearly 80 co-authorships between biological and agricultural faculties. The cooperation on the national and international level is illustrated in Figures 4.15 and 4.16.

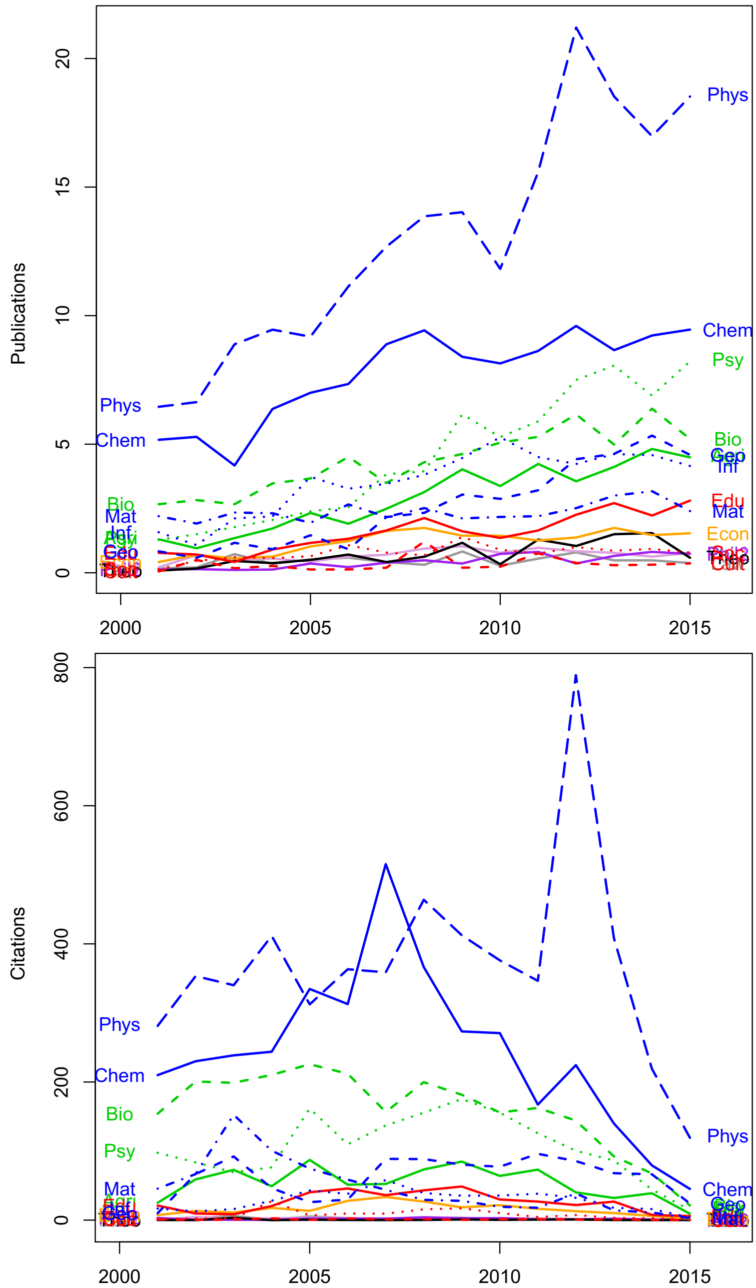

Figure 4.8: Publications (top) and citations count (bottom) per person for faculties from 2001 to 2015 without eight outliers. Citation window equals three years.

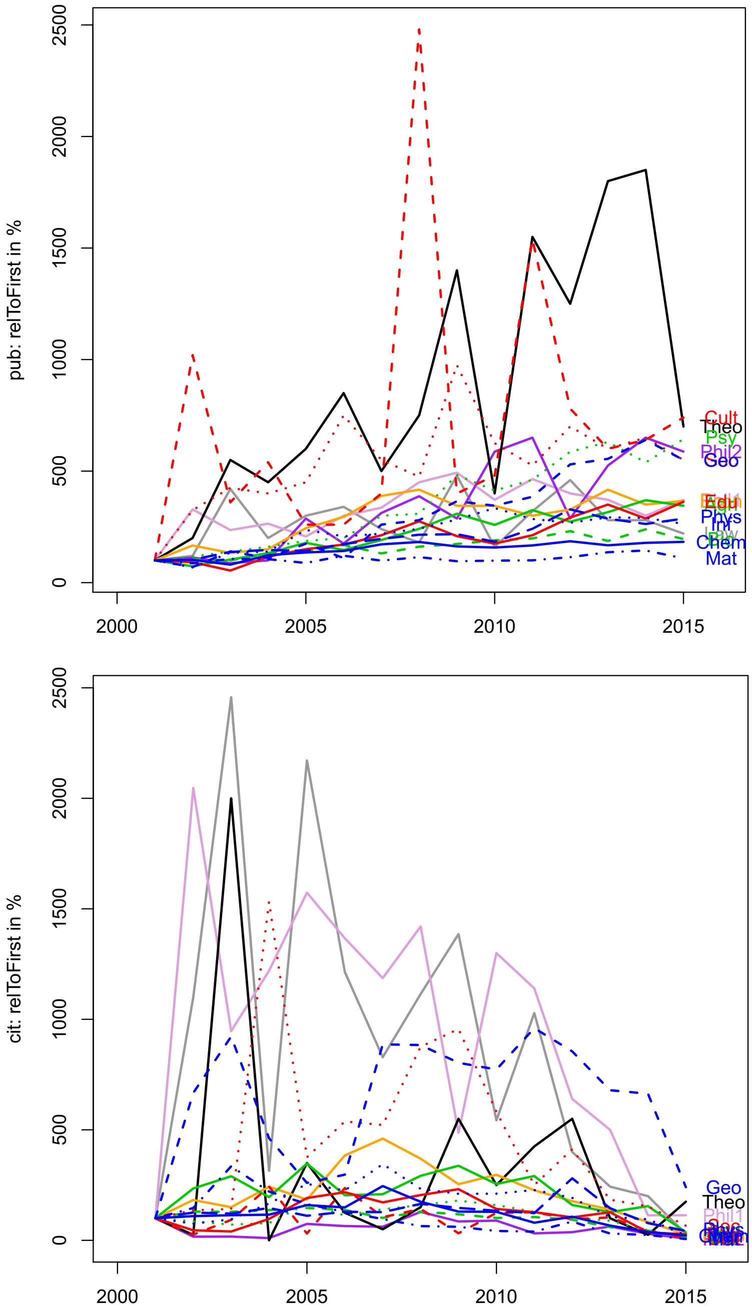

Figure 4.9: Publications (left) and citations (right) growth rate relative to the values 2001 for professorships from 2001 to 2015. The data of eight outliers are removed. Citation window equals three years.

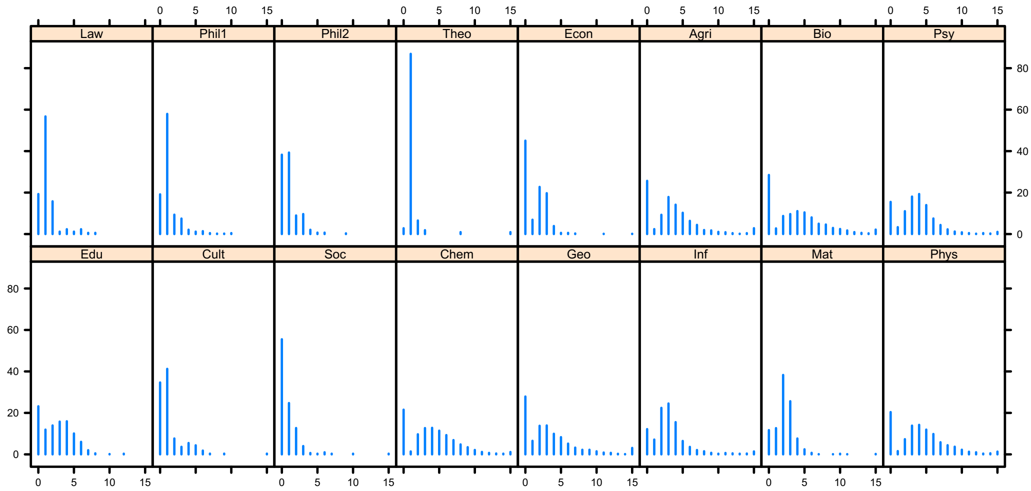

Figure 4.10: Distribution of publications according to the number of authors from 2001 to 2015. The data of eight outliers are removed.

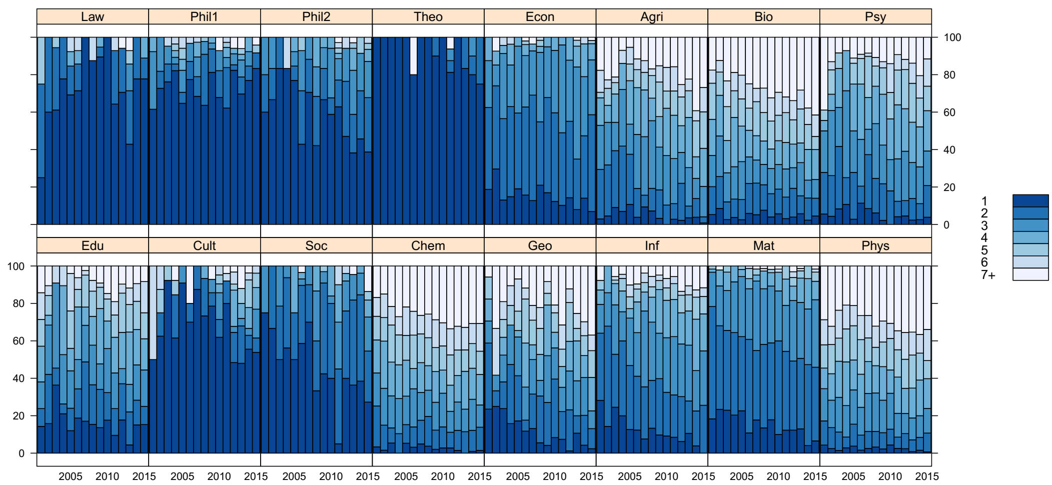

Figure 4.11: Proportion of the number of co-authors (from 1 – dark blue, to >7 – light blue) of publications within faculties. The data of eight outliers are removed.

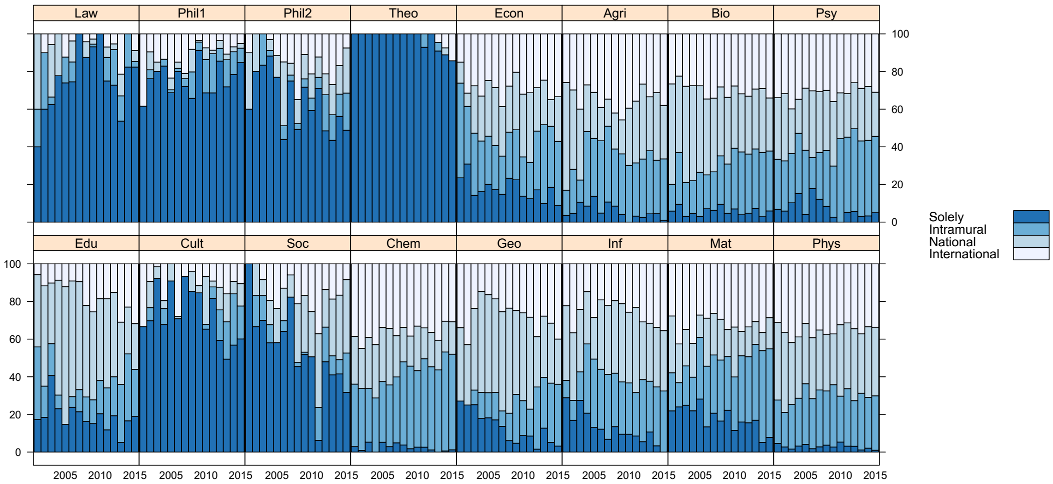

Figure 4.12: Dynamics of cooperation from 2001 to 2015 in percentage: solely authorship (navy blue), multiple inside HU – intramural (dark blue), national (blue) and international (light blue). Fractional counting of publications is used. The data of eight outliers are removed.

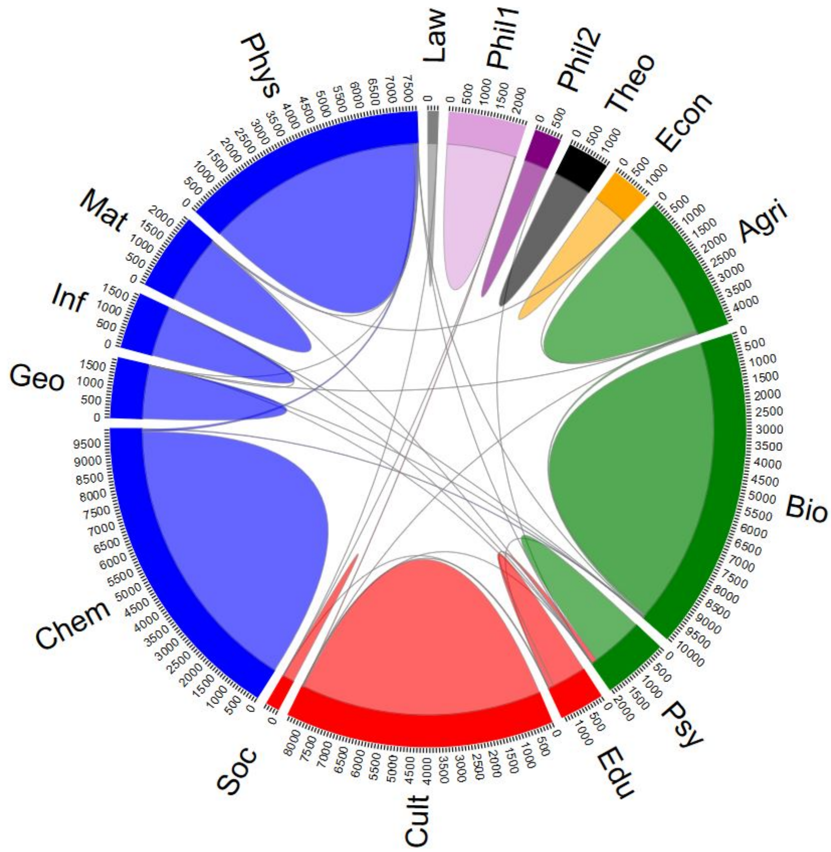

Figure 4.13: Chord diagram for the cooperation within entire university (56579 co-authorships). Full counting, without eight outliers. The color of the outer circle indicates the affiliation to the eight original faculties.

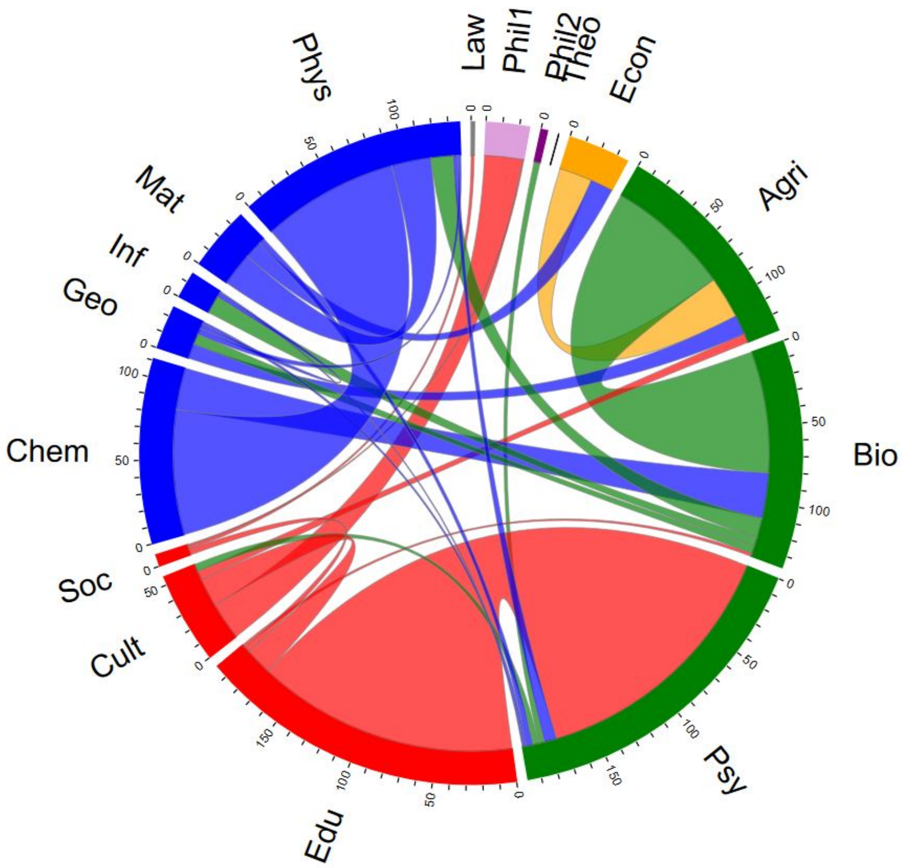

Figure 4.14: Chord diagram for the cooperation within entire university without internal cooperation inside faculties (1122 co-authorships). Full counting, without eight outliers. The color of the outer circle indicates the affiliation to the one of the eight original faculties.

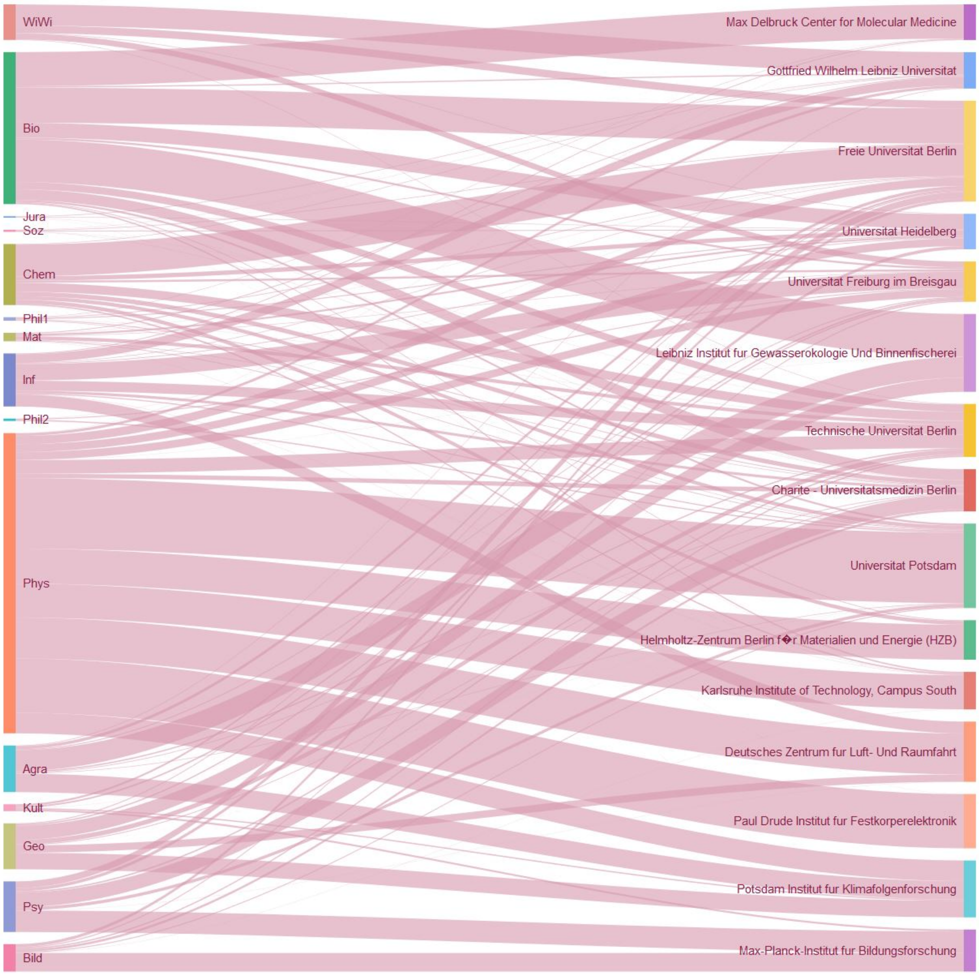

Figure 4.15: National cooperation: Sankey plot for faculties (left) and other German institutions (right), with more than 70 publications, fractional counting. The data of eight outliers are removed.

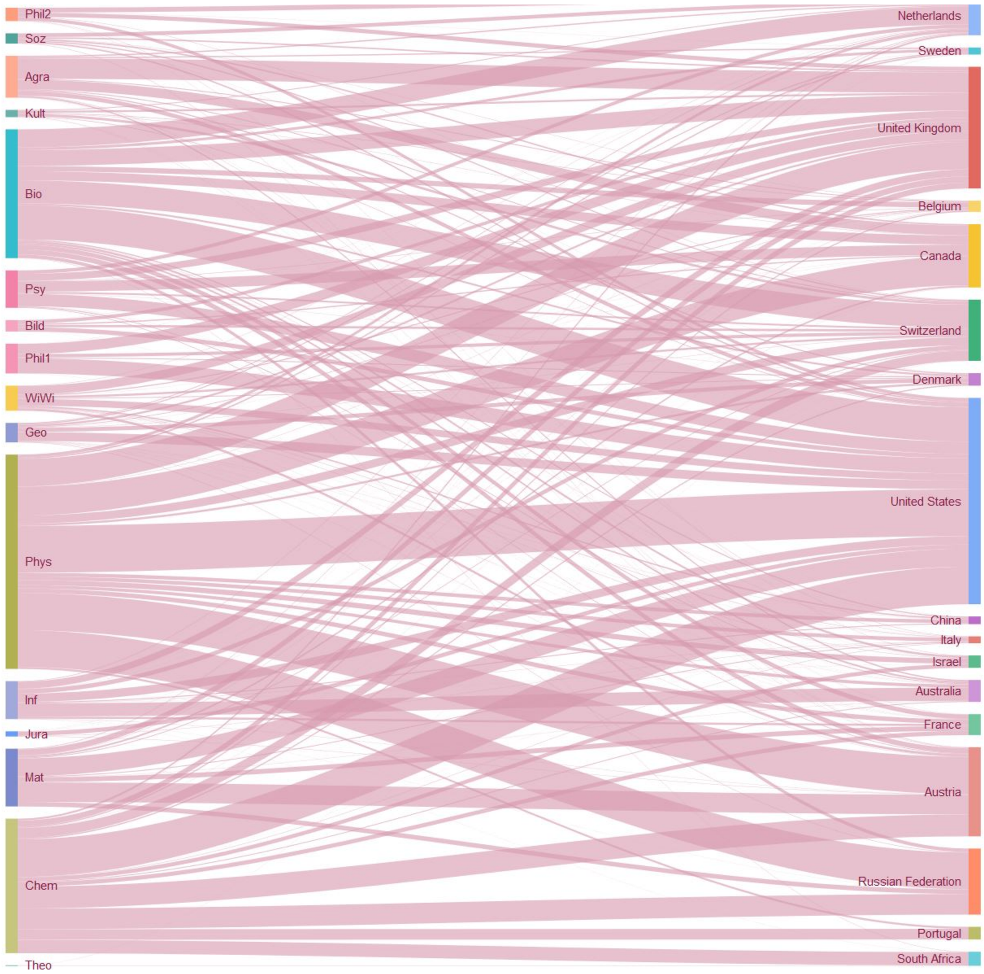

Figure 4.16: International cooperation: Sankey plot for the cooperation between HU units (left) and other countries (right) for 2001–2015, without Germany, fractional counting. The data of eight outliers are removed.

Since similar research fields can be assigned to different faculties, joint publications support the cross-disciplinary character of research. Insight into this cooperation pattern is illustrated on a Sankey plot in Figure 4.17. The left part shows the 16 faculties, whereas the right part introduces 27 research fields taken from Scopus ASJC (All Science Journal Classification). The lines in be-

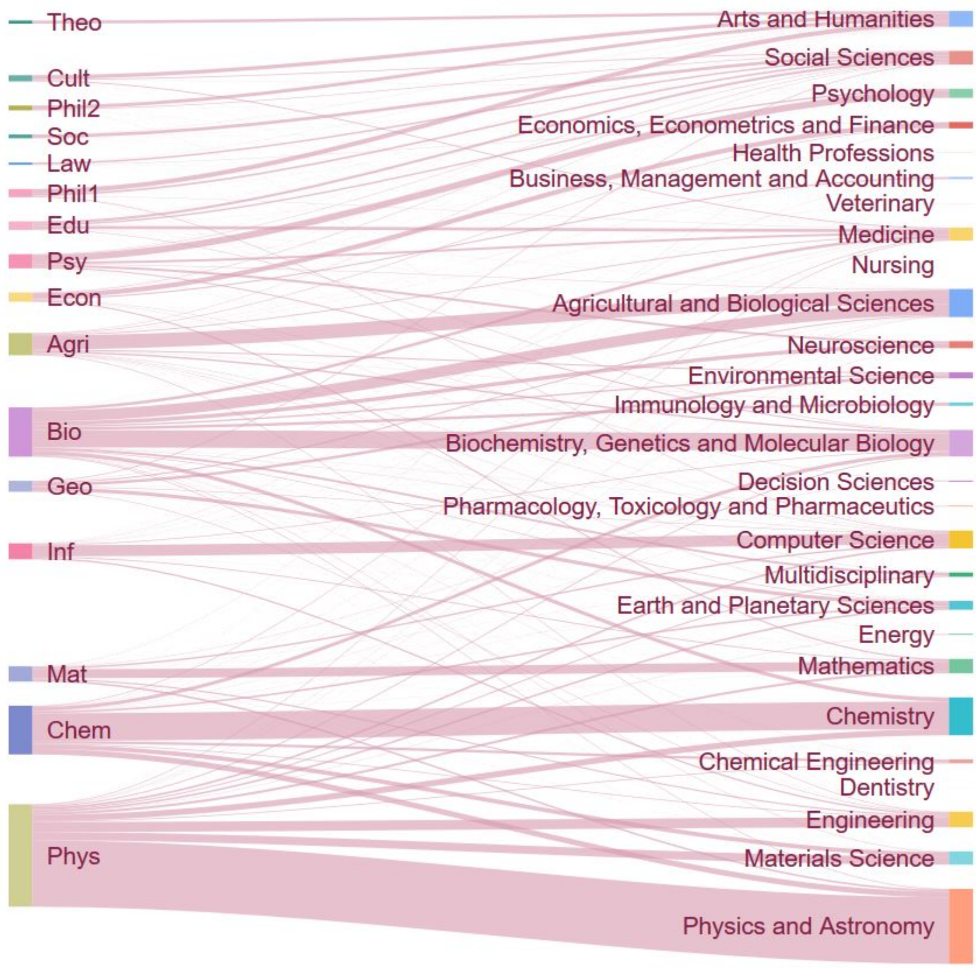

Figure 4.17: Sankey plot for publications published from 2001 to 2015 by professors of eight faculties within 27 research fields. The width of the bars corresponds to the number of publications (28,034 in total). Full counting, without eight outliers.

tween represent the number of publications written by specific faculty members in certain research areas. Each publication is assigned to one of the subject areas according to the main direction of the journal or corresponding outlet where it was published. One conclusion from Figure 4.17 is that the majority of faculty publications is in its primary profiling field. There is, however, a rich set of research outputs arising from other fields. Such an interdisciplinary pattern is evident for mathematical and natural sciences, life sciences, but also economics and educational sciences.

In summary, the collaboration between faculties and, correspondingly, fields predetermines the interdisciplinary structure of research outputs. Not all faculty members cooperate to the same extent with their colleagues from other areas. Therefore, the various analyses between TPE, publications and citations that occur on individual-data-level may capture the heterogeneity and interdisciplinarity of the actual research results and not only the differences of the main field of researchers or their faculty.

4.4 Methodology

4.4.1 PVARX Model

The current state of TPE, publications and citations can be considered as a result of the historical development of each entity (the corresponding autocorrelation functions confirm this time series characteristic). This feature motivates the use of vector autoregressive (VAR) models, which are used in multivariate time series analysis. Since the information on the past is acknowledged additionally to the relationship structure between variables, VAR models allow us to perform the data description, forecasting, structural and policy analysis in a clear and understandable manner (Stock and Watson 2001, Tsay 2014, Pfaff 2008).

The VAR(p) model of order p can be written as (see Lütkepohl 2005, Koop and Korobilis 2010):

$$y_t = \alpha + \sum_{j=1}^{p} A_j y_{t-j} + \varepsilon_t, \;\; \varepsilon_t \sim \mathrm{N}(0, \Sigma), \;\; \text{i.i.d.} \tag{4.1}$$

where $y_t = (y_{1t}, \ldots, y_{Kt})^\top$ is a $(K \times 1)$ vector of observations for $t = 1, \ldots, T$, $\alpha = (\alpha_1, \ldots, \alpha_K)^\top$ is a $(K \times 1)$ vector of intercepts, A_j is a $(K \times K)$ matrix of coefficients, $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Kt})^\top$ is a $(K \times 1)$ vector of errors or innovations and p represents an order. Note that $\mathrm{E}(\varepsilon_t) = 0$, time invariant positive definite covariance matrix $\mathrm{E}(\varepsilon_t \varepsilon_t^\top) = \Sigma_\varepsilon$ and $\mathrm{E}(\varepsilon_t \varepsilon_q^\top) = 0$ for $q \neq t$.

The variables described in (4.1) are interdependent and endogenous. However, when the system has some variables that can affect others, but are not influenced by them, it should be modeled rather through a VAR model with exogenous variables (VARX):

$$y_t = \alpha + \sum_{j=1}^{p} A_j y_{t-j} + \sum_{n=1}^{s} B_n x_{t-n} + \varepsilon_t, \tag{4.2}$$

where $x_t = (x_{1t}, \ldots, x_{Mt})^\top$ is $(M \times 1)$ vector of exogenous variables, B_n is $(K \times M)$ matrix of coefficients, p is order for endogenous variables and s is an order for exogenous variables.

Both models in (4.1) and (4.2) are used for time series observations of a single unit. However, empirical data often deal with multiple units. Such cross-sectional dimension can be handled by the panel VAR (PVAR) model (Canova and Ciccarelli 2013, Dees and Güntner 2014, Abrigo and Love 2016). For real data problems, the PVAR model may appear to be restrictive. This can be avoided by including the exogenous variables into the model (Holtz-Eakin 1988, Juodis 2016, Fomby et al. 2013, Cavallari and D'Addona 2014, Djigbenou-Kre and Park 2016). Consequently, a panel vector autoregressive model with exogenous variables PVARX(p, s) is given by:

$$y_{i,t} = \alpha_i + \sum_{j=1}^{p} A_j y_{i,t-j} + \sum_{n=1}^{s} B_f x_{i,t-n} + \varepsilon_{i,t}. \tag{4.3}$$

We estimate the VARX models using the R package MTS created by Tsay (2015). We also extend the package to support the multiple observations, i.e. to estimate the PVARX model. The R codes are available at GitHub.

4.4.2 Model Specification

Since the hypotheses of interest and the theoretical assumptions regarding the interdependence involve the causal relationship between the variables in both directions, we consider TPE, publications (PUB) and citations (CIT) as dependent variables and specify the model as a system of equations in (4.4). This design allows us to test our hypotheses.

$$
\begin{aligned}
TPE_t &= \alpha_1 + \sum_{i=1}^{p_1} \beta_{1i} TPE_{t-j} + \sum_{i=1}^{m_1} \gamma_{1i} PUB_{t-j} + \sum_{i=1}^{k_1} \delta_{1i} CIT_{t-j} + \zeta_{11} AGE_t + \varepsilon_{1t}, \\
PUB_t &= \alpha_2 + \sum_{i=1}^{m_2} \beta_{2i} TPE_{t-j} + \sum_{i=1}^{p_2} \gamma_{2i} PUB_{t-j} + \sum_{i=1}^{k_2} \delta_{2i} CIT_{t-j} + \zeta_{21} AGE_t + \varepsilon_{2t}, \\
CIT_t &= \alpha_3 + \sum_{i=1}^{m_3} \beta_{3i} TPE_{t-j} + \sum_{i=1}^{k_3} \gamma_{3i} PUB_{t-j} + \sum_{i=1}^{p_3} \delta_{3i} CIT_{t-j} + \zeta_{31} AGE_t + \varepsilon_{3t}.
\end{aligned} \tag{4.4}
$$

For instance, γ_{11} shows the linear dependence of TPE_t on PUB_{t-1} in the presence of TPE_{t-1}, CIT_{t-1} and academic AGE_t. The autoregressive structure of the data is emphasized via the lag operator, in other words, time $t-j$. The academic age might influence other variables (Abramo et al. 2016 and Costas et al. 2010), but because of its nature cannot be affected itself. Therefore, the AGE is reflected in the model as an exogenous variable and is considered only in period t. The forecasting errors are encompassed by the corresponding ε.

To select the order p for the PVARX model, we calculate three information criteria: Akaike (AIC), Bayesian (BIC) and Hannan-Quinn (HQ). As a result of the analysis, all three information criteria choose order $p = 1$ according to their minimal values.

4.5 Empirical Results

4.5.1 Estimation

To understand the interdependence structure of faculties, we estimate the PVARX(1,0) model (4.4) for each faculty separately using least squares (LS) estimation. The results are summarized in Table 4.2, where we do not divide the CuSoEd (cultural, social and education sciences) faculty into three entities due to the lack of data on the lower aggregation level. The estimation results in the last sub-table introduce the average for the whole university.

Table 4.2 introduces the relationships between all variables of interest. For instance, the estimated coefficients in the second column of the Biology Institute (Bio) in sub-table (row 3, panel 1) illustrate how TPE responds to the change in TPE, PUB and CIT in the last period and the current academic AGE after allowing for simultaneous change in other predictors in the provided data. If all other variables are held constant, then for each additional EUR in TPE in the previous year at the Bio faculty, one can expect the current TPE to increase by an average of 88 cents. Further, TPE is predicted to increase by 2,729 EUR given one additional publication in the preceding year. At the same time, CIT is insignificant for TPE. Likewise, after adjusting for simultaneous change in the other predictors, PUB responds with 0.64 publications to the one publication increment in the last year. Generally, TPE and CIT are highly significant for PUB. Nonetheless their effect is minuscule. The intercepts suggest that one expects 5,980 EUR, 1.5 publications and around 82 citations on average for Bio with no TPE, PUB and CIT influence. Interestingly, no statistically significant linear dependence of the mean of TPE on CIT and vice versa is found. The high p-value of the LS estimate of AGE indicates that the academic age is not statistically significant, even at the 10% level for TPE. However, a one year increase of academic AGE decreases the predicted PUB level by 0.01 publications and CIT level by 2.14 citations.

	Law			Phil1			Phil2		
	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t
$const$	1,804.74	0.13***	0.79	5,972.56*	0.39***	1.82***	3,822.44***	0.23***	1.07***
	(1,254.71)	(0.04)	(0.68)	(3,517.76)	(0.02)	(0.06)	(1,165.35)	(0.02)	(0.02)
TPE_{t-1}	0.87***	0.00	0.00	0.91***	0.00***	0.00	0.90***	0.00	0.00**
	(0.01)	(0.00)	(0.00)	(0.03)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)
PUB_{t-1}	1,399.97	0.28***	0.76	-607.89	0.49***	-0.12***	279.35	0.35***	-0.02
	(1,494.93)	(0.05)	(0.81)	(2,549.19)	(0.02)	(0.04)	(1,421.20)	(0.03)	(0.03)
CIT_{t-1}	168.23	0.01	0.03	57.10	0.01***	0.32***	65.79	0.00	0.31***
	(153.80)	(0.00)	(0.08)	(392.73)	(0.00)	(0.01)	(141.61)	(0.00)	(0.00)
AGE_t	291.01***	0.00*	0.01	822.40**	-0.01***	-0.05***	85.74	0.00	-0.04***
	(80.40)	(0.00)	(0.04)	(360.01)	(0.00)	(0.01)	(113.27)	(0.00)	(0.00)

	Theo			Econ			Agri		
	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t
$const$	1,320.18	0.10**	0.12***	-1,558.52	0.18***	4.53***	6,361.74***	0.88***	22.60***
	(1,435.87)	(0.04)	(0.02)	(2,596.29)	(0.05)	(0.91)	(2,220.08)	(0.07)	(3.52)
TPE_{t-1}	0.97***	0.00	0.00*	0.93***	0.00***	0.00	0.98***	0.00***	0.00**
	(0.03)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)
PUB_{t-1}	1,121.83	0.42***	0.01	8,995.35***	0.53***	4.03***	1,169.96**	0.70***	0.94
	(1,537.81)	(0.04)	(0.03)	(1,826.19)	(0.03)	(0.64)	(491.50)	(0.02)	(0.78)
CIT_{t-1}	-218.40	-0.04	-0.06***	-212.44**	0.01***	0.28***	-22.72**	0.00***	0.64***
	(944.46)	(0.03)	(0.02)	(85.50)	(0.00)	(0.03)	(11.54)	(0.00)	(0.02)
AGE_t	113.67	0.02***	0.03***	163.77	0.02***	-0.10	-494.66**	-0.02**	-0.26
	(119.68)	(0.00)	(0.00)	(248.88)	(0.00)	(0.09)	(200.64)	(0.01)	(0.32)

	Bio			Psy			Chem		
	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t		PUB_t	CIT_t
$const$	5,980.67***	1.51***	81.90***	2,882.75	1.58***	66.04***	5,212.58	1.87***	108.70***
	(1,303.04)	(0.06)	(2.28)	(4,948.28)	(0.24)	(2.71)	(6,135.88)	(0.15)	(11.39)
TPE_{t-1}	0.88***	0.00***	0.00	0.79***	0.00***	0.00***	0.87***	0.00***	0.00*
	(0.00)	(0.00)	(0.00)	(0.04)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)
PUB_{t-1}	2,729.29***	0.64***	4.71***	1,807.83**	0.72***	-0.73	4,133.67***	0.74***	-0.68
	(251.07)	(0.01)	(0.44)	(913.43)	(0.04)	(0.50)	(743.76)	(0.02)	(1.38)
CIT_{t-1}	-6.65	0.00***	0.38***	24.75	0.00*	0.31***	-13.71	0.00	0.48***
	(4.84)	(0.00)	(0.01)	(27.64)	(0.00)	(0.02)	(14.08)	(0.00)	(0.03)
AGE_t	-165.12	-0.01**	-2.14***	206.86	-0.04	-2.34***	-595.52**	0.03***	2.76***
	(156.75)	(0.01)	(0.27)	(498.13)	(0.02)	(0.27)	(300.71)	(0.01)	(0.56)

Table 4.2: Estimation results of PVARX(1,0) model. ***, ** and * indicate a statistical significance at 1%, 5% and 10% level, respectively. Standard deviation is provided in brackets. Data: without 8 outliers, TPE are inflation adjusted with the base year 2001, PUB with full counting.

	Geo			Inf			Mat		
	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t
const	579.18	0.85***	23.27***	-769.30	1.16***	11.89***	10,131.11***	0.69***	27.63***
	(3,215.81)	(0.04)	(0.59)	(3,431.72)	(0.12)	(2.06)	(2,570.37)	(0.13)	(3.93)
TPE_{t-1}	0.86***	0.00***	0.00***	0.92***	0.00**	0.00	0.98***	0.00	0.00***
	(0.04)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.02)	(0.00)	(0.00)
PUB_{t-1}	2,323.52**	0.67***	3.33***	3,243.46***	0.69***	3.69***	471.19	0.72***	-0.10
	(916.12)	(0.01)	(0.17)	(841.23)	(0.03)	(0.51)	(719.02)	(0.04)	(1.10)
CIT_{t-1}	8.54	0.00***	0.64***	-49.20	0.00**	0.30***	-4.73	0.00*	0.55***
	(14.84)	(0.00)	(0.00)	(56.68)	(0.00)	(0.03)	(12.22)	(0.00)	(0.02)
AGE_t	429.70	-0.05***	-0.83***	912.11***	-0.02***	-0.50***	-275.78	0.05***	0.00
	(346.87)	(0.00)	(0.06)	(251.73)	(0.01)	(0.15)	(308.02)	(0.02)	(0.47)

	Phys			CuSoEd			University		
	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t	TPE_t	PUB_t	CIT_t
const	10,311.11**	1.60***	107.34***	5.42***	0.22***	1.34***	6,361.40***	0.13***	4.69***
	(5,063.36)	(0.15)	(4.78)	(1.53)	(0.02)	(0.20)	(551.12)	(0.01)	(0.55)
TPE_{t-1}	0.91***	0.00***	0.00	1.00***	0.00***	0.00	0.88***	0.00***	0.00**
	(0.02)	(0.00)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)
PUB_{t-1}	1,349.77***	1.07***	14.36***	1.59	0.51***	2.84***	1,895.95***	0.96***	11.05***
	(373.63)	(0.01)	(0.35)	(1.43)	(0.02)	(0.19)	(181.41)	(0.00)	(0.18)
CIT_{t-1}	-17.16*	0.00***	0.31***	-0.06	0.01***	0.47***	-17.10***	0.00***	0.44***
	(8.83)	(0.00)	(0.01)	(0.07)	(0.00)	(0.01)	(4.58)	(0.00)	(0.00)
AGE_t	2,332.66*	-0.19***	-10.09***	-0.05	0.00	-0.07***	110.97*	0.00	-0.28***
	(1,205.19)	(0.03)	(1.14)	(0.18)	(0.00)	(0.02)	(58.28)	(0.00)	(0.06)

(Table 4.2 continued): Estimation results of PVARX(1,0) model. ***, ** and * indicate a statistical significance at 1%, 5% and 10% level, respectively. Standard deviation is provided in brackets. Data: without 8 outliers, TPE are inflation adjusted with the base year 2001, PUB with full counting.

We are now prepared to check the hypotheses *H1* – *H5*. As shown in Table 4.3, *H1* (TPE drives publications) is rejected for Law, Phil2, Theo and Mat, as the variable TPE_{t-1} is not significant for the variable PUB_t at the $\alpha = 10\%$ significance level. In the same manner, we check the remaining *H2* – *H5* for all faculties and the whole university, see Table 4.3.

Hypotheses	Law	Phil1	Phil2	Theo	Econ	Agri	Bio	Psy	Chem	Geo	Inf	Mat	Phys	CuSoEd	Uni.
H1		+			+	+	+	+	+	+	+		+	+	+
H2					+	+	+	+	+	+	+		+		+
H3			+	+		+		+	+	+		+			+
H4	+	−		+	+	−	−		+	−	−	+	−		
H5	+	+				−			−		+		+		+

Table 4.3: Hypotheses that are rejected (gray) or failed to reject (blue) for each faculty according to the 10% significance level of corresponding variables. The sign denotes the positive (+) or negative (−) influence.

The analysis reveals interesting patterns. First, the social sciences, humanities and mathematics generally have an insignificant relationship between PUB and TPE. Second, the detected influence of TPE on CIT is positive and, interestingly, is present even for fields with no relationship between TPE and PUB. This seems to indicate a difference between quality and quantity of research outputs. Third, the AGE of researchers from the same faculty influences PUB and TPE differently, in the sense of the significance level and sign of the effect. Furthermore, one can clearly see that the results for the whole university in the last column considerably differ from the results of the faculties. This demonstrates that analyses on the high aggregation level of the university do not reflect the behavior of its faculties.

4.5.2 Structural Analysis

Impulse Response Functions

VAR models also provide the possibility to track the reaction of the system given an exogenous impulse. The corresponding impulse response functions (IRF) describe the relations between variables of the system. Orthogonalized IRF allow us to change one variable to the value of its standard deviation shock and to track how the other variable consequently changes over time. The condition is that all other variables have no shocks. Technical details of the methodology can be found in Lütkepohl (1999) and Baltagi (2001). Thus, the IRF shows us how TPE, PUB and CIT change during coming periods, if they are influenced by a specific impulse.

Figure 4.18 shows the dynamic interrelationships within the system from the fitted PVARX(1,0) model with orthogonalized innovations. The first row shows an effect of one standard deviation shock in TPE (panel 1,1), PUB (1,2) and CIT (1,3) on TPE, given there are no other shocks in the system. The second and the third rows introduce the responses in PUB and CIT correspondingly to a specific unit innovation.

The ordering of variables is important for the definition and interpretation of the IRF. We select the order according to the estimation results and implies that TPE is a variable with potential immediate effect on itself and other variables, the shock in PUB can have an instantaneous impact on the last two variables and CIT may influence only the last component of the row. For instance, the first row indicates that TPE may affect all three variables, PUB may influence TPE and CIT, while CIT has potential effect on TPE with some time lag.

The results show that by increasing TPE of Econ in period t_0 by one standard deviation one can expect a 70,000 EUR increase in TPE in the first year t_1, see plot (1,1) in Figure 4.18. In periods t_2 to t_5 from 68,000 EUR to 60,000 EUR of TPE are additionally obtained by Econ on a yearly basis. In other words, Econ may gain approximately 330,000 EUR cumulatively at the end of the fifth year given one standard deviation innovation increase in TPE in the starting period t_0.

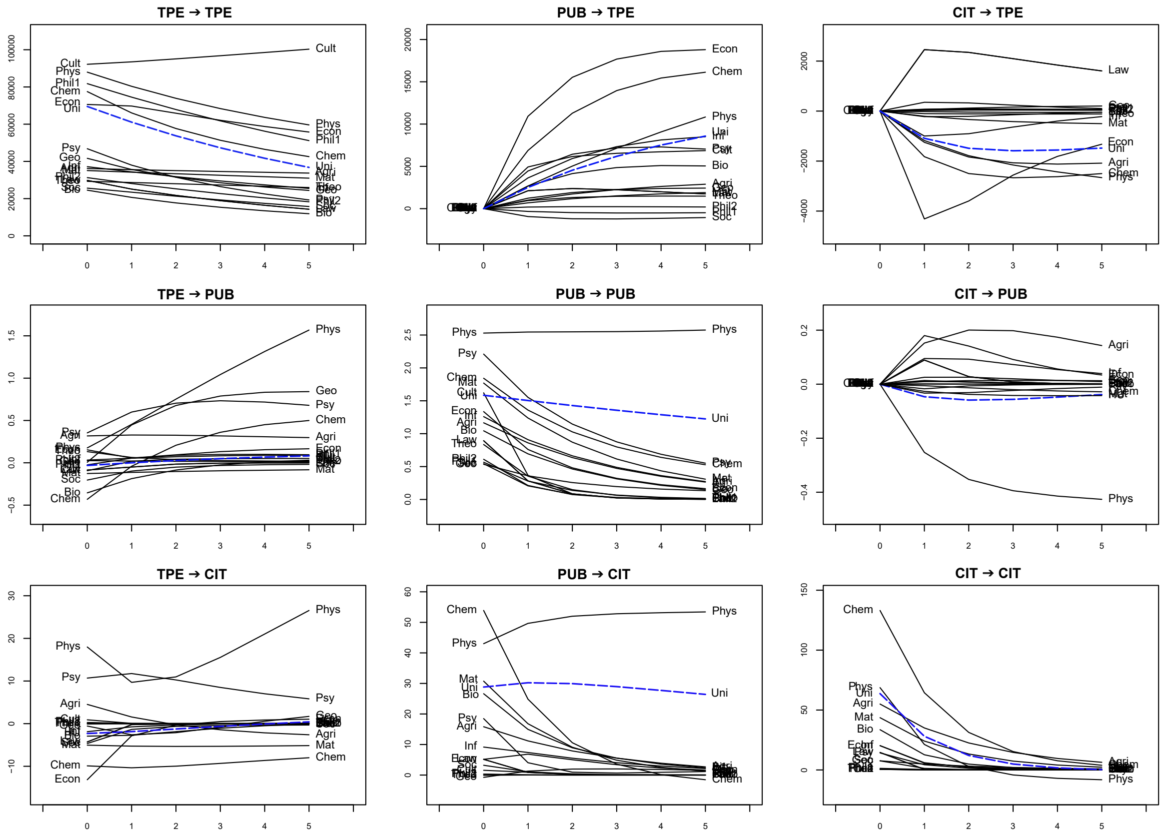

Figure 4.18: Impulse Response Functions of the PVARX(1,0) model for TPE, CIT und PUB for faculties (black lines) and university (blue dashed line) for the first five periods. Innovations are orthogonalized (impulse → response).

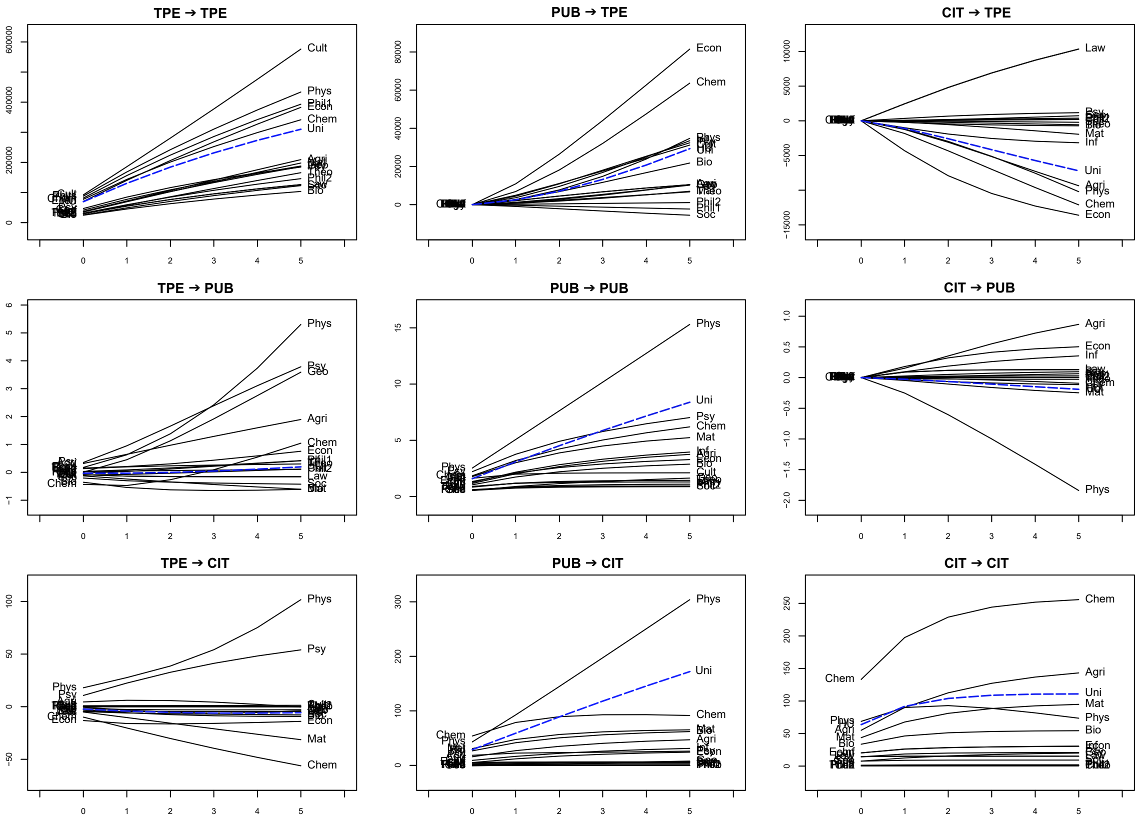

Figure 4.19: Cumulated IRF of the PVARX(1,0) model for TPE, CIT und PUB for faculties (black lines) and university (blue dashed line) for the first five periods. Innovations are orthogonalized (impulse → response).

Similarly, the plot (1,2) shows that the TPE response with a time lag of one period and a value of around 7,000 EUR on an innovation in PUB for Chem. By t_5 the TPE reaches nearly 17,000 EUR per year or nearly 62,000 EUR in total for five periods for Chem. Next, one standard deviation shock in CIT in Econ leads to 400 EUR decrease by the first year and then demonstrates a further gradual decline, as it still stays below zero, see (1,3). A shock in TPE also has an immediate impact on PUB. The plot (2,1) depicts a slow but steady increase of PUB from t_0 to t_5 for almost all faculties when it is influenced by a one standard deviation shock in TPE. A positive but over time declining impact has a shock in PUB on itself for all faculties, except Phys, which is slowly increasing with each additional year, see (2,2). However, an innovation in CIT (2,3) does not lead to considerable changes in PUB in the long-term perspective, again with the exception of Phys. The plot (3,1) shows that CIT remain more or less stable for all faculties but Phys, influenced by a shock in TPE. For CIT a positive effect of PUB (3,2) and CIT (3,3) innovations rapidly dies away during t_1–t_3. Only Phys demonstrates an opposite trend. The response in all variables to a shock in itself is positive, gradually decreasing for TPE over time (Cult being an exception) and more sharply for PUB (Phys being an exception) and CIT in the first three years. The long-term effects or accumulated responses with orthogonal innovations over five periods to a unit shock are introduced in Figure 4.19.

One can see that the impulses in the same variables cause different responses within faculties. However, the IRF of the university introduces an aggregate that diminishes field-specific behavior. For instance, the cumulated IRF for PUB→PUB, CIT→PUB and PUB→CIT show that the Uni-level increase seems to be heavily driven by a single faculty Phys.

To summarize, we see the possible evolution of TPE, PUB and CIT for all faculties and university along a five year time horizon after a shock in t_0 via IRF. Moreover, the IRF results further support the view that university-level decisions, which may affect all heterogeneous faculties, should not be based on university level data.

The issue with IRF is that if some important variables are not included in the system, their effect is captured by innovations and can result in some bias in IRF. The FEVD overcomes this issue as it shows to what extent the change in variables is explained by external shocks.

Forecast Error Variance Decomposition

Forecast Error Variance Decomposition (FEVD) helps to measure the forecast' preciseness of a fitted VAR model (Tsay 2014). It shows which part of the forecast error variance is explained by a shock at a given horizon h. In other words, one expects to see the percentage of the change in the forecast errors of TPE, PUB and CIT relating to the exogenous shocks of these variable. Table 4.4 shows the FEVD results from 1- to 4-step ahead predictions.

The results demonstrate low interrelation between TPE and other time series. For instance, from 96% to 99% of 2- to 4-steps forecast error variance of TPE is accounted for by shocks in TPE. As with the IRF, this can be partly explained by the selected order of variables. Similarly, from 95% to 99% of the error in PUB can be attributed to the innovations in PUB for social sciences, humanities and informatics. For life sciences and geography, the changes in TPE partly explain the variation in PUB. Moving from 1- to 4-steps forecast horizon, the development in the forecast error variance of CIT that can be explained by its own innovations decreases, whereas the contribution of the PUB and in some cases TPE shocks increases. Such slowly growing influence of other variables is also true for TPE and PUB. The variance in CIT for law, natural and life sciences and language sciences is largely explained by variation in PUB and to a smaller extent in TPE. Interestingly, for economics and management sciences the variation in CIT is better explained by errors in TPE than in PUB. Moreover, the error variance in CIT for psychology is accounted for by PUB and TPE innovations to a greater extent than by CIT itself.

Summarizing the FEVD insights, we identify for which faculties TPE, PUB and CIT act as driving forces of the change of the forecast error of corresponding predictions. We find that the variance of the TPE is mainly related to shocks in TPE. A variance change in PUB and CIT partly corresponds to the shocks in all three system's variables for most of the faculties. One can also conclude there are some omitted variables that possibly influence the system. This is especially true for TPE, as innovations in PUB and CIT explain its forecast error variance to a low degree or not at all for different faculties.

The FEVD results also confirm the view that the results for the university do not reflect the features of single faculties. More importantly, one can see the difference in the explained variance for various faculties and correspondingly research fields.

F. error	h	Law TPE	Law PUB	Law CIT	Phil1 TPE	Phil1 PUB	Phil1 CIT	Phil2 TPE	Phil2 PUB	Phil2 CIT	Theo TPE	Theo PUB	Theo CIT	Econ TPE	Econ PUB	Econ CIT
TPE	1	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
	2	99.31	0.29	0.39	99.99	0.00	0.00	99.99	0.00	0.00	99.95	0.05	0.00	98.62	1.19	0.19
	3	98.92	0.50	0.57	99.99	0.00	0.00	99.99	0.00	0.00	99.89	0.11	0.00	97.31	2.47	0.22
	4	98.69	0.64	0.67	99.99	0.00	0.00	99.99	0.00	0.00	99.85	0.15	0.00	96.21	3.59	0.20
PUB	1	0.91	99.09	0.00	0.18	99.82	0.00	0.18	99.82	0.00	2.44	97.56	0.00	1.30	98.70	0.00
	2	1.10	98.01	0.89	0.90	99.08	0.03	0.26	99.74	0.00	2.54	97.42	0.03	1.10	97.56	1.34
	3	1.11	97.92	0.97	2.67	97.29	0.04	0.34	99.66	0.00	2.64	97.32	0.04	1.33	96.72	1.95
	4	1.12	97.91	0.98	4.81	95.15	0.04	0.39	99.61	0.00	2.71	97.26	0.04	1.88	95.96	2.16
CIT	1	7.12	10.26	82.62	0.36	4.30	95.34	16.95	20.37	62.68	1.62	2.71	95.67	28.27	4.34	67.39
	2	7.11	10.50	82.39	0.35	3.94	95.70	16.31	20.19	63.50	1.85	2.81	95.34	25.91	10.47	63.62
	3	7.11	10.52	82.37	0.36	3.93	95.71	16.22	20.15	63.63	2.01	2.81	95.18	24.84	13.48	61.69
	4	7.11	10.52	82.36	0.37	3.96	95.67	16.33	20.11	63.56	2.17	2.81	95.02	24.43	14.79	60.79

F. error	h	Agri TPE	Agri PUB	Agri CIT	Bio TPE	Bio PUB	Bio CIT	Psy TPE	Psy PUB	Psy CIT	Chem TPE	Chem PUB	Chem CIT	Geo TPE	Geo PUB	Geo CIT
TPE	1	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
	2	99.90	0.04	0.06	99.31	0.69	0.00	99.45	0.54	0.00	99.51	0.45	0.03	99.95	0.05	0.00
	3	99.77	0.11	0.13	98.21	1.78	0.01	98.69	1.31	0.00	98.67	1.26	0.07	99.87	0.13	0.00
	4	99.63	0.18	0.18	97.05	2.95	0.01	97.92	2.08	0.01	97.69	2.21	0.10	99.79	0.21	0.00
PUB	1	6.91	93.09	0.00	10.37	89.63	0.00	2.49	97.51	0.00	5.20	94.80	0.00	0.02	99.98	0.00
	2	8.95	90.06	0.99	9.29	90.67	0.04	6.22	93.76	0.01	3.46	96.54	0.00	32.44	67.55	0.01
	3	10.92	86.88	2.20	8.61	91.33	0.07	10.23	89.75	0.02	3.53	96.46	0.00	57.71	42.26	0.02
	4	12.81	84.03	3.16	8.22	91.70	0.08	13.93	86.04	0.02	4.94	95.05	0.01	71.08	28.89	0.03
CIT	1	0.62	7.60	91.78	0.46	38.40	61.13	17.36	51.91	30.73	0.47	14.02	85.51	0.54	0.97	98.50
	2	0.49	8.07	91.44	0.72	41.68	57.60	30.36	43.05	26.59	0.80	13.76	85.44	7.75	2.47	89.78
	3	0.44	8.38	91.18	0.83	43.17	56.00	38.05	38.20	23.75	1.14	13.57	85.30	10.54	5.56	83.90
	4	0.46	8.57	90.98	0.87	43.81	55.32	42.45	35.49	22.06	1.44	13.44	85.11	10.31	8.53	81.16

Table 4.4: Forecast error variance decomposition of the TPE/PUB/CIT system with the forecast horizon h. The color intensity indicates the degree of explained variance (light blue for 1.00%–25.00%, blue for 25.01%–75.00% and darker blue for 75.01%–100%).

		Inf			Mat			Phys			CuSoEd			University		
F. error	h	Variance decomposition			Variance decomposition			Variance decomposition			Variance decomposition			Variance decomposition		
		TPE	*PUB*	*CIT*	*TPE*	*PUB*	*CIT*	*TPE*	*PUB*	*CIT*	*TPE*	*PUB*	*CIT*	*TPE*	*PUB*	*CIT*
TPE	1	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00	100.00	0.00	0.00
	2	99.45	0.51	0.04	99.98	0.02	0.00	99.94	0.05	0.01	99.98	0.01	0.00	99.91	0.07	0.01
	3	99.45	0.51	0.04	99.94	0.05	0.00	99.81	0.16	0.02	99.97	0.02	0.00	99.74	0.23	0.03
	4	97.71	2.24	0.05	99.90	0.09	0.01	99.62	0.34	0.04	99.96	0.03	0.01	99.48	0.47	0.04
PUB	1	0.07	99.93	0.00	0.51	99.49	0.00	0.48	99.52	0.00	0.00	1.00	0.00	0.04	99.96	0.00
	2	0.16	99.46	0.37	0.61	99.34	0.05	1.80	97.72	0.48	0.18	98.56	1.26	0.02	99.96	0.01
	3	0.32	99.07	0.61	0.72	99.18	0.10	3.94	95.13	0.92	0.58	97.18	2.25	0.03	99.94	0.03
	4	0.51	98.76	0.73	0.82	99.03	0.15	6.72	92.06	1.22	1.16	96.08	2.77	0.05	99.91	0.04
CIT	1	0.71	16.84	82.45	0.89	32.88	66.23	4.69	6.84	68.47	1.16	1.83	97.02	0.11	16.97	82.93
	2	0.70	23.45	75.85	1.41	32.60	65.98	4.21	43.57	52.22	1.20	7.69	91.11	0.13	26.41	73.45
	3	0.67	27.01	72.32	2.02	32.33	65.66	4.22	55.11	40.68	1.18	11.41	87.41	0.13	34.53	65.34
	4	0.66	28.82	70.52	2.64	32.07	65.29	4.93	62.13	32.93	1.15	13.28	85.57	0.13	40.88	59.00

(Table 4.4 continued): Forecast error variance decomposition of the TPE/PUB/CIT system with the forecast horizon h. The color intensity indicates the degree of explained variance (light blue for 1.00%–25.00%, blue for 25.01%–75.00% and darker blue for 75.01%–100%).

4.6 Summary and Discussion

The paper contributes to the discussion of interdependence structures between third-party expenses (TPE), publications (PUB) and citations (CIT). Contrary to most previous studies, we use individual data that provides insight at the highest granularity level and leads to more robust results when aggregating to the faculty level.

Analyzing the data, we employ a sophisticated state-of-the-art methodology never before used in the context of research performance. Our work extends the previous research by using a VAR type model that is usually employed in macroeconomic analyses (Holtz-Eakin 1988, Canova and Ciccarelli 2013). The application of the PVARX model on the microeconomic level allows us to capture the interdependencies of multiple time series, take advantage of the cross-sectional dimension and benefit from exogenous variables.

4.6.1 Interpretation of Results

Here we summarize the findings obtained from the PVARX (1,0) model using estimation results, IRF and FEVD in the light of three primary areas: social sciences and humanities (SSH; including Law, Phil1, Phil2, Theo, Econ, Cult, Soc and Educ faculties), life sciences (LSc; Agri, Bio and Psy) and mathematical and natural sciences (MNS; Chem, Geo, Inf, Mat and Phys), see Table 4.1.

Social Sciences and Humanities

We find a positive impact of academic funding on the current number of PUB for Econ, Phil1, Cult, Soc and Educ. However, we identify no such effects for Law, Phil2 and Theo. This may be caused by the research areas of Law, Phil2 and Theo generally attracting fewer TPF than other fields of SSH (more information is available in the Supplemental material). Interestingly, the opposite case, i.e. the past productivity of researchers influences the likeliness of obtaining TPF, is true only for Econ with more effect seen in later periods.

The effect of TPE on CIT is significant only for Phil2 and Theo. This in combination with previous results suggests that Phil2 and Theo may produce research outcomes with higher visibility and acceptance among scientific communities with the same funding as other SSH faculties. Contrary to Payne and Siow (2003), we detect no negative influence of academic funding on the number of CIT for any faculty of the whole university. This may indicate that focusing on obtaining of external funding does not necessarily cause a decrease in the quality of PUB. On the other hand, we find significant negative effect of CIT on TPE for Econ, which continues to decrease gradually over time. This value is also the lowest for the whole university. For instance, one additional CIT in the previous year leads to a decrease in TPE of Econ by around 200 EUR, if all other variables are held constant. To justify this, one may suggest that researchers producing high-quality PUB spend more time for research instead of writing of proposals to attract TPF.

The IRF results show that additional TPE leads to even more TPE in the long-term perspective for all SSH faculties; the corresponding increase for Cult being the largest in the whole university. One can track a similar pattern for PUB. An increase in PUB by one shock increases PUB over next five periods. This is consistent with the FEVD results indicating that innovations in PUB cause the most of the change in PUB.

Our further analyses suggest that the scientific productivity increases with the academic age for Law, Theo and Econ. Although the academic age of researchers in Phil1 leads to a decrease in the number of PUB, it also causes an increase in TPE. This can be referred to the shift of focus over life time or other reasons.

Life Sciences

Our results show a positive, significant impact of TPE on PUB for all LSc faculties. The error variance in PUB is partly accounted for by shocks in TPE for up to 4-step ahead predictions. We also identify the positive influence of PUB on TPE over time and, furthermore, the slow but steady increase of TPE over time given a shock in PUB for LSc.

We find that TPE positively affects CIT for Agri and Psy and causes further sharp increase in CIT for Psy in a 5-year perspective given one additional innovation in TPE. This is consistent with the FEVD results for Psy, showing TPE as a driving force of change in the forecasting error variance in CIT in the long-term perspective. A possible explanation deals with the fact that after receiving a grant, the researcher needs time to carry out experiments, work thoroughly on the research problem and write a research paper. When the research work is published, it starts to collect CIT only after a certain period equal to the length of the citation window.

Interestingly, the variation in PUB explains almost half of the change in CIT for Bio, which is similar to the pattern of Phys from MNS. A possible reason is that some areas of Bio and Phys may have PUB with nearly one hundred of co-authors. As a result, the researchers produce a higher number of PUB, which generate a higher number of CIT. Regarding the academic age, the results are consistent with the previous literature. The age of researchers negatively affects the number of PUB (for Agri and Bio) and TPE (for Agri).

Mathematical and Natural Sciences

Researchers of MNS with more funding produce more PUB and those who publish more attract more TPF. The only exception is Mat, where we discover no significant dependence between TPE and PUB in both directions. A shock in PUB has a positive impact on TPE during the next five years for all MNS faculties. Similarly, the TPE innovations lead to increase in PUB. Furthermore, the highest influence of change in TPE on PUB in the whole university is for Phys. A high proportion of error variance of PUB for Geo is explained by shocks in TPE. This value is also the largest among all faculties. For Chem, innovations in TPE account for the change in PUB to a smaller extent.

External funding has a positive influence on the number of CIT for Chem, Geo and Mat. This is further supported by FEVD for Geo and Mat, as the variance in CIT is explained to a smaller extent by a variation in TPE. The fact that for Mat the TPE cause an increase in CIT but are not significant for the number of PUB suggests that academic funding supports the higher quality of Mat PUB, but not necessarily their quantity.

The academic age of researchers influences PUB and TPE of MNS differently in the sense of both significance level and the sign of the effect. The detected impact of academic age on TPE is positive for Inf and Phys, but negative for Chem. Interestingly, for these faculties the impact of age on PUB has an opposite sign, i.e. negative for Inf and Phys and positive for Chem.

All results show the difference between analyses of the faculties and suggest that performing analyses on the high aggregation level of universities does not reflect the behavior of its faculties.

4.6.2 Implications for Policy and Decision Making

The differences in research fields pose a significant challenge for any policy maker, as the decision influences the whole university. Following our results, the reaction of a single faculty to an exogenous shock may be different from the reaction of other faculties or the effect seen on the aggregated university's level to the same shock. Therefore, the possible consequence of using this university-level information for the setting of incentive mechanisms may be a significant shift in the reacting behavior of researchers.

In the wake of the rise of New Public Management, universities increasingly use research performance measurements for the design of incentive-based motivation. A vivid example is performance-oriented budgeting that, among other targets, aims to stimulate attraction of more TPF and PUB in peer-reviewed journals. The common equal-for-all policy may punish faculties with low need in TPF, publishing mainly in books and with a majority of PUB with a single author – humanities being an example. While areas such as high energy physics may produce less than the world average of the corresponding field, the quantity of the research outputs may be higher than in other fields. Thus, one expects here no additional motivation to produce more, as a result of the performance-oriented policy. Moreover, using the counting of PUB and TPE, which is not field-normalized, to assess the research performance may also have structural effects, such as increasing the number of fragmented PUB, risk aversion and shift of focus from quality to quantity (Butler 2003).

Furthermore, the effects of field diversity may have a serious impact on the governance of a university. In particular, implementing structural reforms (i.e. merger or division of faculties) requires clear understanding of how close the research between fields is; how similar the writing, publishing or citing behavior is; how equivalent the need in TPF is; how intense the cooperation between faculties is; and how strong is the interdisciplinary research involving areas of interest. Providing policy makers with data-driven analyses as provided here (and in the Supplemental material) regarding these issues should complement experts judgments and, as a result, enhance the quality of decisions.

4.6.3 Recommendations for University Research Management

Given increased complexity along with the availability of information that policy and decision makers use for university research management, the questions of how to distinguish the relevant data basis, which methods to use for its analysis, and how to visualize the empirical results in clear and understandable manner are of great importance.

Our findings confirm the significant difference between faculties of a university and corresponding research fields regarding publishing and citing behavior, amount of TPF and practices of their attraction. A comparison of key performance indicators across divisions is common practice for decision making in a managerial environment. In fact, using raw non-adjusted data captures the diversity of the groups. However, it may lead to false conclusions. We emphasize that university management should normalize the research performance indicators for decision making involving comparison across fields. This may help to eliminate the potential effects of research areas and make the performance measurements suitable for the research management process. Whether to perform normalization with the world or national fields' average, depends on the goals of the policy.

A growing need for data-driven support for decision making involves an inextricably linked concern about the reliability of analytical results, which is affected by data quality. Publication and citation datasets, as a rule, originate from external databases (Scopus by Elsevier, Web of Science by Thomson Reuters, Google Scholar, etc.) This creates a bias against disciplines with lesser coverage by bibliometric databases. An important question emerges from this consideration: to what extent can one rely on analyses for a specific faculty or discipline? The possible solution, in the authors' opinion, deals with the establishment of internal bibliometric data management utilizing all published outlets of university members. This, firstly, helps to select the external bibliometric database with the best coverage for the university. Secondly, this provides an evidence about the proportion of covered PUB of researchers and, subsequently, of faculties in a selected database. Thus, the meaningfulness of performance indicators based on such internal database can be justified for each faculty.

Our work deals with information on full professorships and their labs. Including other factors, such as data about teaching, administrative and refereeing duties, into the model may improve its precision. Furthermore, using data of all scientific members of a university (associate professors, assistant professors, research assistants etc.) may lead to including more PUB in the dataset, the possibility to capture more heterogeneity in the model and, as a result, to produce more accurate results.

Universities are a source of knowledge production. Industry benefits from cooperation with universities through the access to the i) knowledge pool; ii) qualified workforce; iii) latest analytical techniques, for instance, econometric methods and data mining. The practice of using scientific methods for the improvement of internal processes at a university itself is often underestimated. Analogous to a business, a university generates a lot of data throughout its activities that represent a rich source of information for decision support. The internal data evaluation using advanced statistical, econometric and data mining techniques available at the research environment of a university is a step towards a better understanding of the current state, explaining the past and making forecasts or describing future trends.

While admitting the critical role of information for the governance of top-level research, the argument about the lack of a workforce to undertake the complex analytical job is still common. One possible remedy is a better use of available resources, i.e. establishing internal research projects involving university scientists or as a part of Bachelor, Master or Ph.D. thesis. The possible data privacy issue should be, of course, accounted for, for example, by working with anonymized, encoded or aggregated data. Such a combination of unique data, vast methodological knowledge and veiled personnel resources, results in a synergy effect for managerial decision making promoting research excellence.

Throughout the paper, we use modern visualization techniques which help to display the complex relationships in an understandable form. Striving to facilitate the cooperation across disciplines and increase the international visibility, research policy makers require targeted informational support. The Sankey plots allow us to understand the interdisciplinary structure of the faculties intuitively. Although not a central aim of this paper, this visualization technique is further applied to check the internationality of the faculties, i.e. with which universities or institutions on the national or international level does every faculty cooperate (see the Supplemental file).

Quantitative analyses provide an important insight into academic collaboration and its productivity. Here, we suggest the use of the chord diagram, a graphical method generally used to display interrelationships of genome data, for mapping of the intramural cooperation structures across faculties. To achieve this, we use joint PUB and information about co-authors to identify and measure inter-faculty channels of cooperation. Equally, one can use research projects and information about principal investigators.

In summary, our results shed light on the complex interdependencies between TPE, PUB and CIT uncovered from individual-level data. The findings from estimation results, IRF and FEVD support the idea that scientific areas have diverse structures. Policy making that affects heterogeneous faculties should account for specifics of individual fields and not only rely on university level indicators. Providing the visualization of sophisticated data facilitates an understanding of the current state and future trends in research performance, helps to sharpen the research profile of the university, and enables a focused approach toward research management. The combination of data-driven analyses with expert knowledge creates significant added value for strategic decision making and further improves the foundations for the successful research management of the university.

A Appendix

A.1 Supplementary materials for Chapter 1

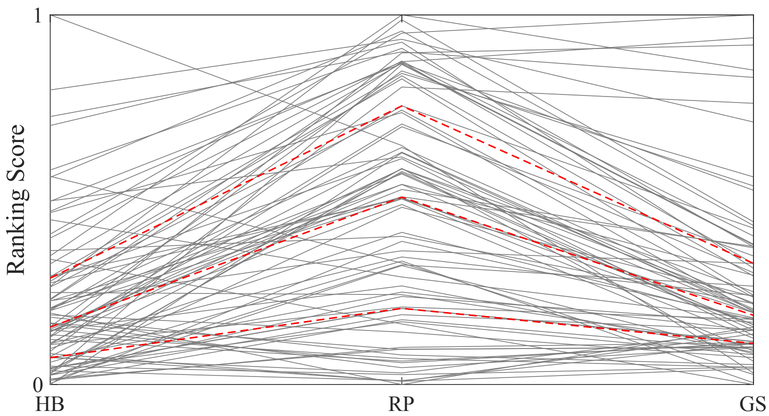

Figure A.1: Parallel coordinate plot for three variables (HB, RP and GS) on 82 researchers. Two outliers from HB and GS are removed. Red lines denote the three quartiles (25%, 50% and 75%). RP values are rescaled.

	Count	Mean	St.dev	Median	Min	Max
HB						
Age	458	47.3	9.5	45.0	29.0	75.0
Common Score	500	7.6	3.8	6.4	4.3	35.8
RP						
Average Rank Score	2304	1107.0	631.7	1100.0	2.8	2194.0
Aabs-Views Score	1435	2640.0	2544.7	1861.0	1052.0	36870.0
Abs-Views Score	1529	4447.0	3494.7	3323.0	1860.0	44760.0
Ad-Cites Score	1922	299.6	304.0	200.4	98.9	3378.0
Adownloads Score	1410	738.6	685.9	520.6	287.0	7766.0
Adsc-Cites Score	1874	852.7	880.0	570.2	244.3	10300.0
Anb-Cites Score	1936	1321.0	1432.8	856.7	404.5	16800.0
Anb-Pages Score	1415	877.2	430.8	754.3	463.5	4486.0
Anb-Works Score	1319	109.3	58.6	92.3	55.8	903.7
Asc-Cites Score	1890	13320.0	15007.4	8274.0	3405.0	162100.0
Asc-Pages Score	1680	13610.0	9677.1	10600.0	5414.0	115800.0
Asc-Works Score	1823	1381.0	1010.8	1046.0	555.8	10210.0
Awdsc-Cites Score	1821	180.0	186.9	118.8	48.7	2081.0
Awsc-Cites Score	1835	685.0	785.3	420.1	162.3	8311.0
Awsc-Pages Score	1614	682.8	500.7	524.3	250.8	5334.0
Awsc-Works Score	1718	79.8	63.1	58.4	28.3	592.5
Between Score	1148	10.8	9.3	7.9	3.6	94.7
Close Score	1223	4.6	0.2	4.6	4.0	4.8
D-Cites Score	1889	500.8	494.8	342.8	162.5	5878.0
Dnb-Works Score	1343	128.5	66.1	111.0	68.0	1091.0
Downloads Score	1444	1273.0	992.3	950.0	511.0	10950.0
Dsc-Cites Score	1840	1444.0	1468.3	956.3	418.9	17640.0
H-Index Score	2017	19.4	7.4	17.0	12.0	78.0
Nb-Cites Score	1951	2113.0	2275.9	1385.0	640.0	29620.0
Nb-Pages Score	1521	1211.0	581.4	1046.0	658.0	6722.0
Nb-Works Score	1456	185.8	94.1	161.0	97.0	1288.0
Ncauthors Score	1898	1113.0	844.1	834.0	425.0	7787.0
Nep-Cites Score	1764	82.1	6.9	82.3	69.2	93.9
Rcauthors Score	1897	854.7	633.8	645.2	326.8	5722.0
Sc-Cites Score	1889	21610.0	24319.3	13500.0	5548.0	313000.0
Sc-Pages Score	1762	19410.0	13171.2	15450.0	8056.0	167500.0
Sc-Works Score	1884	2025.0	1402.1	1567.0	851.8	14870.0
Students Score	1093	814.1	575.2	711.2	4.3	2202.0
Wdsc-Cites Score	1787	306.5	313.3	201.6	83.1	3580.0
Wsc-Cites Score	1834	1114.0	1271.3	697.3	265.4	15220.0
Wsc-Pages Score	1681	980.2	678.3	782.2	377.1	7587.0
Wsc-Works Score	1791	116.9	90.2	87.7	43.8	1007.0
GS						
Total Cites	1438	10190.0	19831.2	5332.0	0.0	234200.0
H Index	1438	32.9	20.2	29.0	0.0	177.0
I Index	1438	66.0	69.4	46.0	0.0	814.0

Table A.1: Descriptive statistics for 42 factors of HB, RP and GS values. Count is the number of observations, mean is the average of values, St.dev - standard deviation, max and min - maximum and minimum values.

	Count	Mean	St.dev	Median	Min	Max
HB						
<36	4	6.3	1.6	5.8	5.0	8.6
36-40	33	5.7	1.5	5.2	4.3	9.8
41-45	97	6.8	2.8	5.9	4.4	22.8
46-50	117	7.2	2.4	6.7	4.4	15.6
51-55	90	7.9	3.9	6.7	4.3	27.1
56-60	53	9.3	4.0	8.1	4.6	22.4
61-65	39	9.4	6.2	6.9	4.4	35.8
66-70	18	10.0	5.3	7.2	5.0	23.6
>70	7	12.2	8.5	9.0	5.0	29.7
RP						
<36	1	341.8	–	341.8	341.8	341.8
36-40	2	372.4	40.8	372.4	343.6	401.3
41-45	15	276.7	117.8	306.1	89.7	473.3
46-50	30	291.2	140.0	304.8	5.2	479.7
51-55	72	291.8	123.0	305.0	2.8	479.5
56-60	94	247.1	142.5	240.1	11.4	487.5
61-65	90	205.7	137.9	184.4	12.7	475.2
66-70	66	219.5	129.3	211.1	9.0	452.8
>70	88	214.8	147.2	189.1	3.4	489.0
GS						
<36	0	–	–	–	–	–
36-40	5	10240.0	1182.5	10840.0	8758.0	11470.0
41-45	26	12600.0	4745.0	11200.0	8075.0	28400.0
46-50	52	12860.0	5179.4	11070.0	7924.0	29670.0
51-55	86	18780.0	22906.8	13460.0	8012.0	212800.0
56-60	101	22640.0	20020.8	14340.0	7932.0	127300.0
61-65	74	25360.0	22591.8	17290.0	8190.0	161000.0
66-70	55	22680.0	17533.9	17740.0	7931.0	92730.0
>70	59	51730.0	61926.0	20680.0	8022.0	234200.0

Table A.2: Descriptive statistics for HB, RP and GS values through age groups indicating the number of observations (count), the average of values (mean), standard deviation (st.dev), maximum (max) and minimum (min) values.

A.2 Supplementary materials for Chapter 3

TPE

Faculty	Per year						Total number
	mean	sd	min	max	skew.	kust.	of unique obs.
Law	75491	134579	-12064	962237	3.02	14	402
Phil1	117628	167645	-11495	1039790	2	11	330
Phil2	59907	83268	-201	413528	2	6	404
Theo	44405	67549	0	398130	2	10	275
Econ	111245	202180	-3828	1514808.01	3	16	414
Agri	164523	182954	0	856494	2	5	204
Bio	223043	338673	-8471	3007184	4	22	288
Psy	139689	110010	0	389293	0	2	81
Edu	98639	108938	0	405850	1	4	123
Cult	98746	223163	-2256	2051762	6	46	479
Soc	87138	127817	-393	563300	2	8	105
Chem	222662	278167	-258	2173382	3	16	219
Geo	95888	118846	0.00	581977	2	8	93
Inf	153125	222632	0	1171008	3	10	147
Mat	139733	158113	0	678933	1	4	161
Phys	276034	322817	-2099	2291122	32	13	219

Table A.3: Descriptive statistics for third-party funds.

PUB

Faculty	Per year						Total number
	mean	sd	min	max	skew.	kust.	of unique obs.
Law	1.56	1.23	1	9	3.17	15.35	131
Phil1	1.79	1.50	1	11	2.53	10.39	359
Phil2	1.60	1.05	1	9	3.25	18.48	231
Theo	1.80	0.99	1	5	1.31	4.57	144
Econ	2.55	2.14	1	14	2.37	9.48	307
Agri	4.65	4.04	1	19	1.26	3.92	377
Bio	5.07	4.66	1	38	3.25	18.99	684
Psy	5.34	4.20	1	21	1.29	4.62	184
Edu	3.79	3.29	1	17	1.22	3.97	140
Cult	1.43	0.86	1	8	4.08	26.04	329
Soc	1.99	1.36	1	9	1.98	8.12	138
Chem	8.44	6.03	1	31	1.02	3.70	432
Geo	4.14	4.25	1	24	2.15	8.20	191
Inf	4.73	3.50	1	16	1.21	3.90	200
Mat	3.59	3.21	1	16	1.69	5.50	261
Phys	16.02	21.33	1	166	3.85	20.80	488

Table A.4: Descriptive statistics for publications.

CIT

Faculty	Per year						Total number
	mean	sd	min	max	skew.	kust.	of unique obs.
Law	7.02	16.91	0	152	5.77	45.09	131
Phil1	5.59	11.63	0	110	4.89	34.17	359
Phil2	7.00	15.49	0	98	3.50	16.25	231
Theo	1.35	2.48	0	11	2.95	11.52	144
Econ	33.41	51.95	0	347	3.17	14.98	307
Agri	89.62	184.21	0	1390	3.88	20.99	377
Bio	188.84	238.39	0	1710	2.79	12.98	684
Psy	117.36	133.34	1	1052	2.85	16.35	184
Edu	52.77	77.96	0	294	1.75	4.99	140
Cult	4.62	19.69	0	231	8.20	81.82	329
Soc	23.14	56.18	0	521	5.88	47.63	138
Chem	265.86	305.05	0	2639	2.82	15.58	417
Geo	112.17	232.74	0	1706	4.39	25.51	191
Inf	40.06	58.23	0	307	2.18	8.24	200
Mat	81.98	192.23	0	1642	5.36	38.21	261
Phys	477.81	831.19	0	8284	5.40	45.26	488

Table A.5: Descriptive statistics for citations.

Bibliography

Abramo, G., D'Angelo C.A. and Di Costa, F. 2009. Research collaboration and productivity: is there correlation? *Higher Education* **57**(2) 155–171.

Abramo, G. and D'Angelo, C.A. 2011. National-scale research performance assessment at the individual level. *Scientometrics* **86** 347–364.

Abramo, G., D'Angelo, C.A. and Murgia, G. 2016. The combined effects of age and seniority on research performance of full professors. *Science and Public Policy* **43**(3) 301–319.

Abramo, G., Cicero, T. and D'Angelo, C.A. 2011. Assessing the varying level of impact measurement accuracy as a function of the citation window length. *Journal of Informetrics* **5**(4) 659–667.

Abramo, G., Cicero, T. and D'Angelo, C.A. 2013. Individual research performance: A proposal for comparing apples to oranges. *Journal of Informetrics* **7**(2) 528–539.

Abrigo, M.R.M. and Love, I. 2016. Estimation of panel vector autoregression in Stata. *Stata Journal* **16**(3) 778–804.

Baltagi, B. 2001. *A Companion to Theoretical Econometrics* Blackwell, Oxford. ISBN 978-0-631-21254-6.

Barra, C. and Zotti, R. 2016. Measuring Efficiency in Higher Education: An Empirical Study Using a Bootstrapped Data Envelopment Analysis. *International Advances in Economic Research* **22** 11-33.

Bartol, T., Budimir, G., Dekleva-Smrekar, D., Pusnik, M., and Juznic, P. 2014. Assessment of research fields in Scopus and Web of Science in the view of national research evaluation in Slovenia. *Scientometrics* **98**(2) 1491–1504.

Baum, C.F. 2013. Quantile Regression. *Lecture notes on Applied Econometrics, Boston College.*

Beaudry, C. and Allaoui, S. 2012. Impact of public and private research funding on scientific production: The case of nanotechnology. *Research Policy* **41**(9) 1589–1606.

Bergman, L.E.M. 2012. Finding Citations to Social Work Literature: The Relative Benefits of Using Web of Science, Scopus, or Google Scholar. *University Libraries Faculty Scholarship* **18**.

Birks, Y., Fairhurst, C., Bloor, K., Campbell, M., Baird, W. and Torgerson, D. 2014. Use of the h-index to measure the quality of the output of health services researchers. *Journal of Health Services Research & Policy* **19**(2) 102–109.

Bolli, T. and Somogyi, F. 2011. Do competitively acquired funds induce universities to increase productivity? *Research Policy* **40**(1) 136–147.

Bonaccorsi, A. and Daraio, C. 2003. Age effects in scientific productivity. *Scientometrics* **58**(1) 49–90.

Borke, L. and Härdle, W.K. 2017. GitHub API based QuantNet Mining infrastructure in R. *SFB 649 Discussion Paper* **2017**(008).

Borke, L. and Härdle, W.K. 2018. Q3-D3-LSA, *in Handbook of Big data Analytics*, Härdle, W.K., Lu, H. and Shen, X. (Eds.) Springer Verlag. ISBN 978-3-319-18284-1.

Bornmann, L. 2013. How to analyze percentile citation impact data meaningfully in bibliometrics: The statistical analysis of distributions, percentile rank classes, and top-cited papers. *Journal of the American Society for Information Science and Technology* **64**(3) 587–595.

Boyack, K.W. and Börner, K. 2003. Indicator-assisted evaluation and funding of research: Visualizing the influence of grants on the number and citation counts of research papers. *Journal of the American Society for Information Science and Technology* **54**(5) 447–461.

Butler, L. 2003. Explaining Australiaâs increased share of ISI publicationsâthe effects of a funding formula based on publication counts. *Research Policy* **32**(1) 143–155.

Butz, A. and Wohlrabe, K. 2016. Die Ökonomen-Rankings 2015 von Handelsblatt, FAZ und RePEc: Methodik, Ergebnisse, Kritik und Vergleich. *Ifo Working Paper Series* **212**.

Cameron, A. C. and Miller, D.L. 2015. A Practitioner's Guide to Cluster-Robust Inference. *Journal of Human Resources* **50**(2) 317–373.

Canova, F. and Ciccarelli, M. 2013. Panel Vector Autoregressive Models: A Survey. *C.E.P.R. Discussion Papers* **2013**(9380).

Carayol, N. and Matt, M. 2004. Does research organization influence academic production? Laboratory level evidence from a large European university. *Research Policy* **33**(8) 1081–1102.

Cavallari, L. and D'Addona, S. 2014. Trade margins and exchange rate regimes: new evidence from a Panel VARX model. *CREI Working Paper* **2014**(5).

Cole, S. 1979. Age and scientific performance. *American Journal of Sociology* **84**(4) 958–977.

Combes, P.-P. and Linnemer, L. 2010. Inferring Missing Citations - A Quantitative Multi-Criteria Ranking of all Journals in Economics. *HAL Working Papers* **halshs-00520325**.

Costas, R., van Leeuwen, T.N. and Bordons, M. 2010. A bibliometric classificatory approach for the study and assessment of research performance at the individual level: The effects of age on productivity and impact. *Journal of the American Society for Information Science and Technology* **61**(8) 1564–1581.

CRC (Collaborative Research Center) 649 "Economic Risk". `http://sfb649.wiwi.hu-berlin.de/about/index.php` accessed 27 Aug 2017.

Dees, S. and Güntner, J. 2014. Analysing and forecasting price dynamics across euro area countries and sectors: A panel VAR approach. *ECB Working Paper Series* **2014**(1724).

DFG (Deutsche Forschungsgemeinschaft). 2015. Förderatlas 2015. Kennzahlen zur öffentlich finanzierten Forschung in Deutschland. *Wiley-VCH Verlag GmbH & Co.* KGaA, Weinheim, Germany.

DFG (Deutsche Forschungsgemeinschaft; eng. – German Research Foundation). `www.dfg.de/en//research_funding/programmes/coordinated_programmes/collaborative_research_centres/index.html` accessed 2 January 2018.

Diem, A. and Wolter, S.C. 2013. The use of bibliometrics to measure research performance in educational sciences. *Research in Higher Education* **54** 86–114.

Dilger, A. and Müller, H. 2011. Ein Forschungsleistungsranking auf der Grundlage von Google Score. *Diskussionpapier des Instituts für Organisationsökonomik* **2011**.

Dilger, A. and Müller, H. 2013. A citation-based ranking of German-speaking researchers in business administration with data of Google Scholar. *European Journal of Higher Education* **3**(2) 140–150.

Djigbenou-Kre, M.-L. and Park, H. 2016. The effects of global liquidity on global imbalances. *International Review of Economics and Finance* **42**(C) 1–12.

Donner, P. and Aman, V. 2015. Quantilbasierte Indikatoren für impact und Publikationsstartegie. Ergebnisse für Deutschland in allen Fachdisziplinen in den Jahren 2000 bis 2011. *Studien zum deutschen Innovationssystem* IFQ, **8**.

Dyckhoff, H., Rassenhövel, S. and Sandfort, K. 2009. Empirische Produktionsfunktion betriebswirtschaftlicher Forschung: Eine Analyse der Daten des Centrums für Hochschulentwicklung. *Schmalenbachs Zeitschrift für betriebswirtschaftliche Forschung* **61**(1) 22–56.

Ebadi, A. and Schiffauerova, A. 2015. Bibliometric Analysis of the Impact of Funding on Scientific Development of Researchers. Conference Proceedings. *International Conference on Information Systems and Technology Management (ICISTM)*, Montreal, Canada.

Ebadi, A. and Schiffauerova, A. 2016. How to boost scientific production? A statistical analysis of research funding and other influencing factors. *Scientometrics* **106**(3) 1093–1116.

Fomby, T., Ikeda, Y. and Loayza, N. 2013. The growth aftermath of natural disasters. *Journal of Applied Econometrics* **28**(3) 412–434.

Forschungsmonitoring Portal. *Verein für Socialpolitik*. www.forschungsmonitoring.org (accessed 02 January 2018).

Gerhards, J. 2013. Der deutsche Sonderweg in der Messung von Forschungsleistungen. *Wissenschaftspolitik im Dialog*. Berlin-Brandenburgische Akademie der Wissenschaften **7**.

German Council of Science and Humanities (Wissenschaftsrat). 2011. *Recommendations on the Assessment and Management of Research Performance*. Halle Drs. **1656–11**.

GitHub. Programming codes to the doctoral thesis. `www.github.com/AlonaZharova` (accessed 2 January 2018).

Glänzel, W. and Schubert, A. 2003. A new classification scheme of science fields and subfields designed for scientometric evaluation purposes. *Scientometrics* **56**(3) 357–367.

Grözinger, G. and Leusing, B. 2006. Wissenschaftsindikatoren an Hochschulen. *Discussion Papers*, Europa-UniversitÃt Flensburg, International Institute of Management. **012**.

Hamermesh, D.S. 2015. Citations in Economics: Measurement, Uses and Impacts. *NBER Working Paper* **Nov 2015**(21754).

Harzing, A.K. and Wal, R. 2008. Google Scholar as a new source for citation analysis. *Ethics in science and environmental politics* **8** 61–73.

Hausman, J., Hall, B.H. and Griliches, Z. 1984. Econometric Models for Count Data with an Application to the Patents-R & D Relationship. *Econometrica* **52**(4) 909–938.

Hicks, D., Wouters. P., Waltman, L., de Rijcke, S. and Rafols, I. 2015. Bibliometrics: The Leiden Manifesto for research metrics. *Nature* **520**(7548) 429–431.

Hirsch, J.E. 2005. An index to quantify an individual?s scientific research output. *Proceedings of the National Academy of Sciences of the United States of America* **102**(46) 16569–16572.

Holtz-Eakin, C.W.J. 1988. Estimating Vector Autoregressions with Panel Data. *Econometrica* **56**(6) 1371–1395.

Hood, C. 1991. A public management for all seasons? *Public Administration* **69**(1) 3–19.

Hornbostel, S. 1991. Drittmitteleinverbung. Ein Indikator für universitäre Forschungsleistungen. *Beiträge zu Hochschulforschung* **1** 57–84.

Hornbostel, S. 2001. Third party funding of German universities. An indicator of research activity? *Scientometrics* **50**(3) 523–537.

Hornbostel, S. and Torger, M. 2015. Die Exzellenzinitiative und das deutsche Wissenschaftssystem: Eine bibliometrische Wirkungsanalyse. *Wissenschaftspolitik im Dialog*. Berlin-Brandenburgische Akademie der Wissenschaften **12**.

Jacob, B.A. and Lefgren, L. 2011. The impact of research grant funding on scientific productivity. *Journal of Public Economics* **95**(9) 1168–1177.

Jansen, D., Wald, A., Franke, K., Schmoch, U. and Schubert, T. 2007. Drittmittel als Performanzindikator der Wissenschaftlichen Forschung. *Koelner Z.Soziol.u.Soz.Psychol* **59**(1) 125–149.

JEL (Journal of Economic Literature) Classification System. `www.aeaweb.org/econlit/jelCodes.php?view=jel` (accessed 02 January 2018)

Juodis, A. 2016. First difference transformation in panel VAR models: Robustness, estimation, and inference. *Econometric Reviews* 1–44.

Kelchtermans, S. and Veugelers, R.L. 2011. The great divide in scientific productivity: why the average scientist does not exist. *Industrial and Corporate Change* **20**(1) 295–336.

Koop, G. and Korobilis, D. 2010. Bayesian Multivariate Time Series Methods for Empirical Macroeconomics. *Foundations and Trends in Econometrics* **3**(4) 267–358.

Koenker, R. and Bassett, G. 1978. Regression Quantiles. *Econometrica* **46**(1) 33–50.

Koenker, R. and Hallock, K.F. 2001. Quantile Regression. *Journal of Economic Perspectives* **15**(4) 143–156.

Koenker, R. 2005. *Quantile Regression*. Econometric Society Monograph, Cambridge University Press, Cambridge, ISBN 0-521-84573-4, ISBN 0-521-60827-9

Koenker, R. 2015. Quantile Regression in R: A Vignette. 2015. `www.semanticscholar.org/paper/Quantile-Regression-in-R-a-Vignette-Koenker/09e7c24de82e7e081470110ea086deea25113eb6` (accessed 2 Januar 2018).

Kyvik, S. 1990. Age and scientific productivity. Differences between fields of learning. *Higher Education* **19**(1) 37–55.

Larivière, V., Macaluso, B., Archambault, È. and Gingras, Y. 2010. Which scientific elites? On the concentration of research funds, publications and citations. *Research Evaluation* **19**(1) 45–53.

Laudel, G. 2005. Is external research funding a valid indicator for research performance? *Research Evaluation* **14**(1) 27–34.

Lee, G.J. 2010. Assessing publication performance of research units: extensions through operational research and economic techniques. *Scientometrics* **84**(3) 717–734.

Levin, S.G. and Stephan, P.E. 1991. Research Productivity Over the Life Cycle: Evidence for Academic Scientists. *The American Economic Review* **81** 114–132.

Lütkepohl, H. 2005. *New introduction to multiple time series analysis*. Springer Berlin, ISBN 3-540-40172-5.

Lütkepohl, H. 1999. Vector autoregressive analysis. *SFB 373 Discussion Papers* **1999**(31).

McAllister, P.R. and Wagner, D.A. 1981. Relationship between R&D expenditures and publication output for U.S. colleges and universities. *Research in Higher Education* **15**(1) 3–30.

Moed, H.F. 2005. *Citation analysis in research evaluation*. Springer Dordrecht, Great Britain. ISBN 978-1-4020-3714-6.

Moed, H.F. 2010. Measuring contextual citation impact of scientific journals. *Journal of Informetrics* **4**(3) 265–277.

Moed, H.F., de Moya-Anegón, F., López-Illescas, C. and Visser, M. 2011. Is concentration of university research associated with better research performance? *Journal of Informetrics* **5** 649–658.

Mongeon, P., Brodeur, C., Beaudry, C. and Lariviere, V. 2016. Concentration of research funding leads to decreasing marginal returns. *Research Evaluation* **25**(4) 396–404.

Nag, S., Yang, H., Buccola, S. and Ervin, D. 2013. Productivity and financial support in academic bioscience. *Applied Economics* **45**(19) 2817–2826.

NRC (National Research Council). 2010. A Data-Based Assessment of Research-Doctorate Programs in the United States. 2010. The National Academies Press.

Oberschelp, A. and Jaeger, M. 2015. Leistungsvergleiche als Instrument der Hochschulsteuerung: Ansätze, organisatorischer Kontext und Unterstützung des Steuerungshandelns. *Bibliotheksdienst* **49**(5) 475–494.

Pastor, J.M. and Serrano, L. 2016. The determinants of the research output of universities: specialization, quality and inefficiencies. *Scientometrics* **1029**(2) 1255–1281.

Pastor, J.M., Serrano, L. and Zaera, I. 2015. The research output of European higher education institutions. *Scientometrics* **102**(3) 1867–1893.

Payne, A. and Siow, A. 2003. Does Federal Research Funding Increase University Research Output? *The B.E. Journal of Economic Analysis & Policy* **3**(1) 1–24.

Pfaff, B. 2008. VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software* **27**(4) 1–32.

QuantNet. `www.quantlet.de` (accessed 2 January 2018)

Rauber, M. and Ursprung, H.W. 2008. Life Cycle and Cohort Productivity in Economic Research: The Case of Germany. *German Economic Review* **9**(4) 431–456.

Research Excellence Framework. 2011. Assessment framework and guidance on submissions. *REF 02.2011* July 2011.

RePEc (Research Papers in Economics). `www.repec.org` (accessed 02 January 2018).

Rosenbloom, J.L., Ginther, D.K., Juhl, T. and Heppert, J.A. 2015. The Effects of Research and Development Funding on Scientific Productivity: Academic Chemistry, 1990-2009. *PLoS ONE* **10**(9) 1–23.

Schläpfer, F. and Schneider, F. 2010. Messung der akademischen Forschungsleistung in den Wirtschaftswissenschaften: Reputation vs. Zitierhaufigkeiten. *Perspektiven der Wirtschaftspolitik* **11**(4) 325–339.

Schläpfer, F. 2011. Reformbedarf bei der Rating-Agentur für Ökonomen. *Neue Züricher Zeitung* **198** 1–23. `www.felixschlaepfer.ch/Oekonomen-Ranking_26.8.11.pdf` (accessed 2 Januar 2018).

Schmoch, U. and Schubert, T. 2009. Sustainability of incentives for excellent research – The German case. *Scientometrics* **81**(1) 195–218.

Schröder, S., Welter, F., Leisten, I., Richert, A. and Jeschke, S. 2014. Research performance and evaluation?Empirical results from collaborative research centers and clusters of excellence in Germany. *Research Evaluation* **23**(3) 221–232.

Sims, C. 1980. Macroeconomics and Reality. *Econometrica* **48**(1) 1–48.

Sousa, R. 2008. Research Funding: Less Should Be More. *Science* **322**(5906) 1324–1325.

Stegehuis, C., Litvak, N. and Waltman, L. 2015. Predicting the long-term citation impact of recent publications. *Journal of Informetrics* **9**(3) 642–657.

Stock, J.H. and Watson, M.W. 2001. Vector Autoregressions. *Journal of Economic Perspectives* **15**(4) 101–115.

Stock, J.H. and Watson, M.W. 2003. *Introduction to Econometrics*. 1 ed. Pearson, ISBN 0-201-71595-3.

THE (Times Higher Education). Best universities in Germany 2018. *Times Higher Education's World University Rankings data*. `www.timeshighereducation.com/student/best-universities/best-universities-germany` (accessed 2 January 2018).

Tsay, R.S. 2014. *Multivariate Time Series Analysis: with R and Financial Applications*. Wiley. ISBN 978-1-118-61790-8.

Tsay, R.S. 2015. Package "MTS". All-Purpose Toolkit for Analyzing Multivariate Time Series (MTS) and Estimating Multivariate Volatility Models. `https://CRAN.R-project.org/package=MTS` (accessed 2 January 2018).

Van den Berghe, H., Houben, J.A, de Bruin, R.E., Moed, H.F., Kint A., Luwel, M. and Spruyt E.H.J. 1998. Bibliometric indicators of university research performance in Flanders. *Journal of the American Society for Information Science* **49**(1) 59–67.

van Leeuwen, T.N., Moed, H.F., Tijssen, R.J.W., Visser, M.S. and van Raan, A.F.J. 2001. Language biases in the coverage of the science citation index and its consequences for international comparisons of national research performance. *Scientometrics* **51**(1) 335–346.

van Raan, A.F.J., van Leeuwen, T.N. and Visser, M.S. 2011. Severe language effect in university rankings: particularly Germany and France are wronged in citation-based rankings. *Scientometrics* **88**(2) 495–498.

Waltman, L. 2015. Field-normalized citation impact indicators and the choice of an appropriate counting method. *Journal of Infometrics* **9**(4) 872–894.

Waltman, L. 2016. A review of the literature on citation impact indicators. *Journal of Infometrics* **10**(2) 365–391.

Wang, J. 2013. Citation time window choice for research impact evaluation. *Scientometrics* **94**(3) 851–872.

Wildgaard, L. 2016. A critical cluster analysis of 44 indicators of author-level performance. *Journal of Informetrics* **10**(4) 1055–1078.

Wissenschaftsrat. 2004. *Empfehlungen zu Rankings im Wissenschaftssystem Teil 1: Forschung*, Hamburg **Drs. 6285-04**.

Wissenschaftsrat. 2011. *Empfehlungen zur Bewertung und Steuerung von Forschungsleistung* **Drs. 1656–11**.

Wissenschaftsrat. 2012. *Bericht der Steuerungsgruppe zur Pilotstudie zur Weiterentwicklung des Forschungsratings*, Köln **Drs. 2815–12**.

Wohlrabe, K. 2011. Das Handelsblatt- und das RePEc-Ranking im Vergleich. *ifo Schnelldienst* **17** 66–71.

Wohlrabe, K. 2013. Einige Anmerkungen zum Handelsblatt-Ranking 2013. *ifo Schnelldienst* **23** 79–83.

Wooldridge, J.M. 1999. Distribution-Free Estimation of Some Nonlinear Panel Data Models. *Journal of Econometrics* **90** 77–97.

Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. 1 ed. MIT Press Books.

Wooldridge, J.M. 2016. *Introductory Econometrics: A Modern Approach*. 6 ed. Cengage Learning.

Zheng, J. and Liu, N. 2015. Mapping of important international academic awards. *Scientometrics* **104**(3) 763–791.

Zimmermann, C. 2013. Academic Rankings with RePEc. *ifo Econometrics* **1**(3) 249–280.

Selbständigkeitserklärung

Ich erkläre, dass ich die vorliegende Arbeit selbständig und nur unter Verwendung der angegebenen Literatur und Hilfsmittel angefertigt habe.

Berlin, den 8. Januar 2018 Alona Zharova